

SIGNAL PEPTIDE

HUMAN	-19	MQIELSTCFF	LCLLRFCFS
PIG		MQLELSTCVF	LCLLPLGFS
MOUSE		MQIALFACFF	LSLFNFCSS
		** * * * *	*

FIG. 1A

A1 DOMAIN

HUMAN	1	ATRRYYLGAV	ELSWDYMQSD	LG-ELPVADAR	FPPRVPKSFP	FNTSVVYKKT
PIG		AIRRYYLGAV	ELSWDYRQSE	LLRELHVDTR	FPATAPGALP	LGPSVLYKKT
MOUSE		AIRRYYLGAV	ELSWNYIQSD	LLSVLHTDSR	FLPRMSTSFP	FNTSIMYKKT
		*****	*****	*	*	*****
	50	LFVEFTDHLF	NIAKPRPPWM	GLLGPTIQAE	VYDTVVITLK	NMASHPVSLH
		VFVEFTDQLF	SVARPRPPWM	GLLGPTIQAE	VYDTVVVTLK	NMASHPVSLH
		VFVEYKDQLF	NIAKPRPPWM	GLLGPTIWTE	VHDTVVITLK	NMASHPVSLH
		***	***	*****	*****	*****
	100	AVGVSYWKAS	EGAEYDDQTS	QREKEDDKVF	PGGSHTYVWQ	VLKENGPMAS
		AVGVFSWKSS	EGAEYEDHTS	QREKEDDKVL	PGKSQTYVWQ	VLKENGPTAS
		AVGVSYWKAS	EGDEYEDQTS	QMEKEDDKVF	PGESHTYVWQ	VLKENGPMAS
		*****	*****	*****	*****	*****
	150	DPLCLTYSYL	SHVDLVKDLN	SGLIGALLVC	REGSLAKEKT	QTLHKFILLF
		DPPCLTYSYL	SHVDLVKDLN	SGLIGALLVC	REGSLTRERT	QNLHEFVLLF
		DPPCLTYSYM	SHVDLVKDLN	SGLIGALLVC	KEGSLSKERT	QMLYQFVLLF
		*****	*****	*****	*****	*****
	200	AVFDEGKSWH	SETKNSLMQD	RDAASARAWP	KMHTVNGYVN	RSLPGLIGCH
		AVFDEGKSWH	SARNDWSWTRA	MDPAPARAQP	AMHTVNGYVN	RSLPGLIGCH
		AVFDEGKSWH	SETNDSYTQS	MDSASARDWP	KMHTVNGYVN	RSLPGLIGCH
		*****	*	*****	*****	*****
	250	RKSVYWHVIG	MGTTPEVHSI	FLEGHTFLVR	NHRQASLEIS	PITFLTAQTL
		KKSVYWHVIG	MGTSPEVHSI	FLEGHTFLVR	HHRQASLEIS	PLTFLTAQTF
		RKSVYWHVIG	MGTTPEIHSI	FLEGHTFFVR	NHRQASLEIS	PITFLTAQTL
		*****	***	*****	*****	*****
	300	LMDLGQFLLF	CHISSHQHDG	MEAYVKVDSC	PEEPQLRMKN	NEEAEDYDDD
		LMDLGQFLLF	CHISSHHHGG	MEAHVRVES	AEEPQLRKA	DE-EEDYDDN
		LIDLGQFLLF	CHISSHKHDG	MEAYVKVDSC	PEESQWQKKN	NN-EEMEDYD
		*****	*****	*****	*****	*****
	350	LTDSEMDVVR	FDDDNNSPSFI	QIR		
		LYDSDMDVVR	LDGDDVSPFI	QIR		
		DDLYSEMDMF	TLDYDSSPFI	QIR		
			**	***		

IIa/Xa

FIG. 1B

A2 DOMAIN

HUMAN	373	SVAKKHPKTW	VHYIAAEEEED	WDYAPLVLAP	DDRSYKSQYL	NNGPQRIGRK
PIG		SVAKKHPKTW	VHYISAAEEEED	WDYAPAVPSP	SDRSYKSLYL	NSGPQRIGRK
MOUSE		SVAKKYPKTW	IHYISAAEEEED	WDYAPSVPPTS	DNGSYKSQYL	SNGPHRIGRK
		*****	***	*****	*****	***

423	YKKVRFMAYT	DETFKTREAI	QHESGILGPL	LYGEVGDTLL	IIFKNQASRP
	YKKARFVAYT	DVTFKTRKAI	PYESGILGPL	LYGEVGDTLL	IIFKNKASRP
	YKKVRFIAYT	DETFKTRETI	QHESGLLGPL	LYGEVGDTLL	IIFKNQASRP
	***	***	***	*****	*****

A2 INHIBITOR EPITOPE

473	YNIYPHGIDT	VRPLYSRRLP	KGVKHLKDFP	ILPGEIFKYK	WTVTVEDGPT
	YNIYPHGIDT	VSALHPGRLL	KGWKHLKDMP	ILPGETFKYK	WTVTVEDGPT
	YNIYPHGIDT	VSPLHARRLP	RGIKHVKDLP	IHPGEIFKYK	WTVTVEDGPT
	*****	***	***	***	*****

F. IXa BINDING
 APC

523	KSDPRCLTRY	YSSFVNMERD	LASGLIGPLL	ICYKESVDQR	GNOIMSDKRN
	KSDPRCLTRY	YSSSINLEKD	LASGLIGPLL	ICYKESVDQR	GNQMMMSDKRN
	KSDPRCLTRY	YSSFINPERD	LASGLIGPLL	ICYKESVDQR	GNQMMMSDKRN
	*****	***	***	*****	***

573	VILFSVFDEN	RSWYLTENIQ	RFLPNPAGVQ	LEDPEFQASN	IMHSINGYVF
	VILFSVFDEN	QSWYLAENIQ	RFLPNPDGLQ	PQDPEFQASN	IMHSINGYVF
	VILFSIFDEN	QSWYITENMQ	RFLPNAAKTQ	PQDPGFQASN	IMHSINGYVF
	*****	***	***	***	*****

623	DSLQLSVCLH	EVAYWYILSI	GAQTDLSVF	FSGYTFKHKM	VYEDTLTLFP
	DSLQLSVCLH	EVAYWYILSV	GAQTDLSVF	FSGYTFKHKM	VYEDTLTLFP
	DSLELTVCCLH	EVAYWHILSV	GAQTDLSIF	FSGYTFKHKM	VYEDTLTLFP
	***	***	***	*****	*****

673	FSGETVFMMSM	ENPGLWILGC	HNSDFRNRRGM	TALLKVSSCD	KNTGDYYEDS
	FSGETVFMMSM	ENPGLWVLGC	HNSDLRNRRGM	TALLKVYSCD	RDIGDYYDNT
	FSGETVFMMSM	ENPGLWVLGC	HNSDFRKRRGM	TALLKVSSCD	KSTSDYYEEI
	*****	*****	*****	*****	***

◆ IIa/Xa/APC

723	YEDISAYLLS	KNNAIEPR			
	YEDIPIGFLLS	GKNVIEPR			
	YEDIPTQLVN	ENNVIDPR			
	****	*	***		

FIG. 1C

FIG. 1D-1

1287	IVEKYACTTR	ISPNTSQQNF	VTQRSKRALK	QFRLPLEETE	LEKRIIVDDT
	-----	-----	-----	-----	-----
	-GPGRSA	VPRRVKQSLK	QIRLPLEEIK	PERGVVLNAT	
	MVKNYP-----	SQKNI	TTQRSKRALG	QFRL-----	
1337	STQWSKNMKH	LTPSTLTQID	YNEKEKGAI	QSPLSDCLTR	SHSIPQANRS
	STRWS-----	-----	-----	-----	-----
	STQWLKTINC	STQCIIKQID	HSKEMKKFIT	KSSLSDS-SV	IKSTTQTNNS
	***	**			
1387	PLPIAKVSSF	PSIRPIYLTR	VLFQDNSSHL	PAASY----R	KKDGVQESS
	-----	-----	-----	-----	ESS
	DSHIVKTSAF	P---PIDLKR	SPFQNKFHV	QASSYIYDFK	TKSSRIQESN
					**
1433	HFLQGAKKNN	LSLAILTLEM	TGDQREVGSL	GTSATNSVTY	KKVENTVLPK
	PILQGAKRNN	LSLPFLTLEM	AGGQGKISAL	GKSAAGPLAS	GKLEKAVLSS
	NFLKETKINN	PSLAILPWNM	FIDQGKFTSP	GKSNTNSVTY	KKRENIIFLK
	*	*	*	*	*
1483	PDLPKTSGKV	ELLPKVHIYQ	KDLFPTETSN	GSPGHLDLVE	GSLLOGTEGA
	AGLSEASGKA	EFLPKVRVHR	EDLLPQKTSN	VSCAHGDLGQ	EIFLQKTRGP
	PTLPEESGKI	ELLPQVSIQE	EEILPTETSH	GSPGHLNLMK	EVFLQKIQGP
	***	*	***	*	***
1533	IKWNEANRPG	KVPFLRVATE	SSAKTPSKLL	DPLAWDNHYG	TQIPKEEWKS
	VNLNKVNRPG	-----	-----	-----G	PPMPKE-WES
	TKWNKAKRHG	ESIKGKTES-	-SKNTRSKLL	NHHAWDYHYA	AQIPKDMWKS
	*	*	*	*****	*
1583	QEKSPEKTAF	KKKDTI-LSLN	ACESNHAI	INEGQNKP	EVTWAKOQRT
	LEKSPKSTAL	RTKDIISLPLD	RHESNHSIAA	KNEGQAETQR	EAAWTKQGGP
	KEKSPEIISI	KQEDTI-LSLR	PHGNHSIGA	-NEKQNWPQR	ETTWVKQGQT
	*****	*	*	***	***

FIG. 1D-2

FIG. 1E

A3 DOMAIN

FIG. 1F

C1 DOMAIN

HUMAN	2020	KCQTPLGMAS	GHIRDFOITA	SGOYGOWAPK	LARLHYSGSI	NAWSTKEPFS
PIG		ECQAPLGMAS	GRIRDFOITA	SGQYGQWAPK	LARLHYSGSI	NAWSTKDPHS
MOUSE		QCQIPLGMAS	GSIRDFQITA	SGHYGQWAPN	LARLHYSGSI	NAWSTKEPFS
		*****	*****	*****	*****	*****
	2070	WIKV DLLAPM	IIHGIKTQGA	RQKFSSLYIS	QFIIMYSLDG	KKWQTYRGNS
		WIKV DLLAPM	IIHGIIMTQGA	RQKFSSLYIS	QFIIMYSLDG	RNWQSYRGNS
		WIKV DLLAPM	IVHGIKTQGA	RQKFSSLYIS	QFIIMYSLDG	KKWLSYQGNS
		*****	***	*****	*****	***
	2120	TGTL MVFFGN	VDSSGIKHNI	FN PPIIARYI	RLHPTHYSIR	STLRMELMGCDLN
		TGTL MVFFGN	VDASGIKHNI	FN PPIVARYI	RLHPTHYSIR	STLRMELMGCDLN
		TGTL MVFFGN	VDSSGIKHS	FN PPIIARYI	RLHPTHSSIR	STLRMELMGCDLN
		*****	***	*****	*****	*****

FIG. 1G

C2 DOMAIN

INHIBITOR EPITOPE						
HUMAN	2173	SCSMPLGMES	KAISDAQITA	SSYFTNMFAT	WSPSKARLHL	QGRSNAWRPQ
PIG		SCSMPLGMQ	KAISDSQITA	SSHLSNIFAT	WSPSQARLHL	QGRTNAWRPR
MOUSE		SCSIPLGMES	KVISDTQITA	SSYFTNMFAT	WSPSQARLHL	QGRTNAWRPQ
		*****	***	*****	***	*****
	2223	VNNPKEWLQV	DFQKTMKVTG	VTTQGVKSLL	TSMYVKEFLI	SSSQDGHQWT
		VSSAEEWLQV	DLQKTVKVTG	ITTQGVKSLL	SSMYVKEFLV	SSSQDGRRWT
		VNDPKQWLQV	DLQKTMKVTG	IITQGVKSLF	TMSFVKEFLI	SSSQDGHHWT
		*	*****	*****	*****	*****
					C2	
	2273	LFFQNGKVKV	FQGNQDSFTP	VVNSLDPPLL	TRYLRIHPOS	WVHQIALRME
		LFLQDGHTKV	FQGNQDSSTP	VVNALDPPLF	TRYLRIHPTS	WAQHIALRLE
		QILYNGKVKV	FQGNQDSSTP	MMNSLDPPLL	TRYLRIHPQI	WEHQIALRLE
		*	***	*****	*****	***
					PHOSPHOLIPID	
	2323	<u>BINDING</u>				
		VLGCEAQDLY				
		VLGCEAQDLY				
		ILGCEAQQQY				
		*****	*			

FIG. 1H

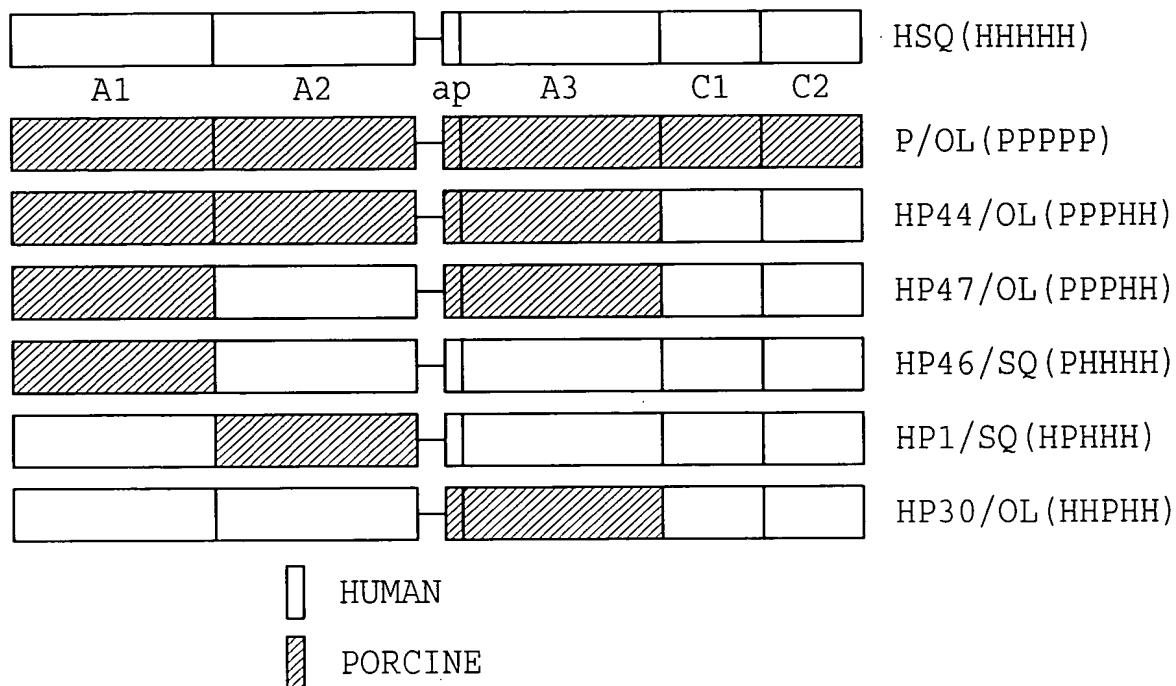


FIG. 2

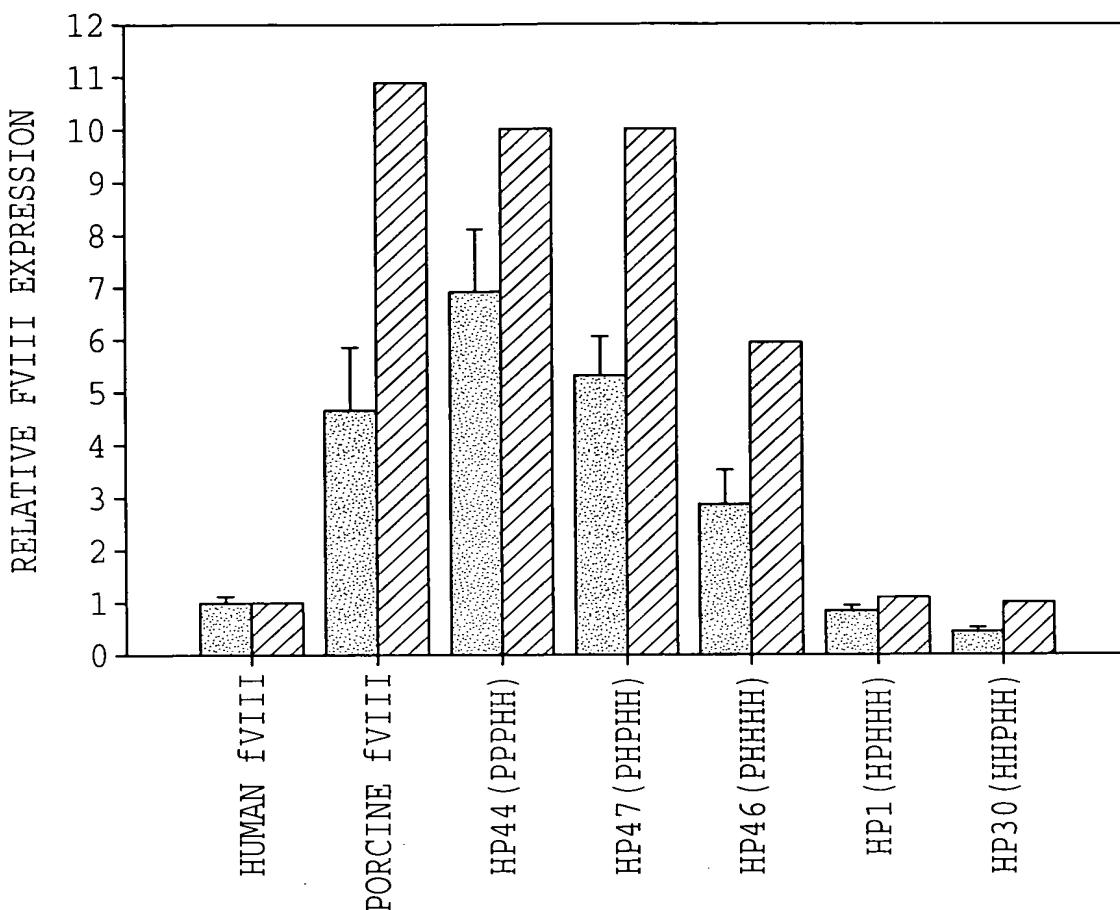


FIG. 3

AMINO ACID SEQUENCE OF HP44/OL

1	MQLELSTCVF	LCLLPLGFSA	IRYYLGAVE	LSWDYRQSEL	LRELHVDTRF
51	PATAPGALPL	GPSVLYKKTV	FVEFTDQLFS	VARPRPPWMG	LLGPTIQAEV
101	YDTVVVTLKN	MASHPVSLHA	VGVSFWKSSE	GAEYEDHTSQ	REKEDDKVLP
151	GKSQTYWQV	LKENGPTASD	PPCLTYSYLS	HVDLVKDLNS	GLIGALLVCR
201	EGSLTRERTQ	NLHEFVLLFA	VFDEGKSWHS	ARNDSWTRAM	DPAPARAQPA
251	MHTVNGYVNR	SLPGLIGCHK	KSVYWHVIGM	GTSPEVHSIF	LEGHTFLVRH
301	HRQASLEISP	LTFLTAQTFL	MDLGQFLLFC	HISHHHHGGM	EAHVRVESCA
351	EEPQLRRKAD	EEEDYDDNLY	DSDMDVVRLD	GDDVSPFIQI	RSVAKKHPKT
401	WVHYIISAEEE	DWDYAPAVPS	PSDRSYKSLY	LNSGPQRIGR	KYKKARFVAY
451	TDVTFKTRKA	IPYESGILGP	LLYGEVGDTL	LIIIFKNKASR	PYNIYPHGIS
501	DVSALHPGRL	LKGWKHLKDM	PILPGETFKY	KWTVTVEDGP	TKSDPRCLTR
551	YYSSSINLEK	DLASGLIGPL	LICYKESVDO	RGNQMMMSDKR	NVILFSVFDE
601	NQSWYLAENI	QRFLPNPDGL	QPQDPFQAS	NIMHSINGYV	FDSLQLSVCL
651	HEVAYWYIILS	VGAQTDLSV	FFSGYTFKHK	MVYEDTLTLF	PFSGETVFMS
701	MENPGLWVLG	CHNSDLRNRG	MTALLKVYSC	DRDIGDYYDN	TYEDIPIGFL
751	SGKNVIEPRS	FAQNSRPPSA	SAPKPPVLRR	HQRDISLPTF	QPEEDKMDYD
801	DIFSTETKGE	DFDIYGEDEN	QDPRSFQKRT	RHYFIAAVEQ	LWDYGMSESP
851	RALRNRAQNG	EVPRFKVVF	REFADGSFTQ	PSYRGELNKH	LGLLGPYIRA
901	EVEDNIMVTF	KNQASRPYSF	YSSLISYPDD	QEQQAEPRHN	FVQPNETRTY
951	FWKVQHHMAP	TEDEFDCKAW	AWFSDVDLEK	DVHSGLIGPL	LICRANTLNA
1001	AHGRQVTVQE	FALFFTIFDE	TKSWYFTENV	ERNCRAPCHL	QMEDPTLK
1051	YRFHAINGYV	MDTLPGGLVMA	QNQRIRWYLL	SMGSNENIHS	IHFSGHVFSV
1101	RKKEEYKMAV	YNLYPGVFET	VEMLPSKVGI	WRIECLIGEH	LOAGMSTTFL
1151	VYSKKCQTPL	GMASGHIRDF	QITASQYGO	WAPKLARLHY	SGSINAWSTK
1201	EPFSWIKVDL	LAPMIIHGIK	TQGARQKFSS	LYISQFIIMY	SLDGKKWQTY
1251	RGNSTGTLMV	FFGNVDSSGI	KHNI FNPPII	ARYIIRLHPTH	YSIRSTLRME
1301	LMGCDLNSCS	MPLGMESKAI	SDAQITASSY	FTNMFATWSP	SKARLHQGR
1351	SNAWRPQVNN	PKEWLQVDFQ	KTMKVTGVTT	QGVKSLLTSM	YVKEFLISS
1401	QDGHQWTLFF	QNGKVKVFQG	NQDSFTPVNN	SLDPPLLTRY	LRIHPQSWVH
1451	QIALRMEVLG	CEAQDLY*			

1-19 SIGNAL PEPTIDE
 20-391 A1 DOMAIN
 392-759 A2 DOMAIN
 760-783 OL LINKER
 784-1154 ap-A3
 1155-1307 C1 DOMAIN
 1308-1467 C2 DOMAIN

FIG. 4

10/33

HP44/OL NUCLEOTIDE SEQUENCE

1 ATGCAGCTAG AGCTCTCCAC CTGTGTCTTT CTGTGTCTCT TGCCACTCGG
TACGTCGATC TCGAGAGGTG GACACAGAAA GACACAGAGA ACGGTGAGCC
51 CTTTAGTGCC ATCAGGAGAT ACTACCTGGG CGCAGTGGAA CTGTCTGGG
GAAATCACGG TAGTCCTCTA TGATGGACCC GCGTCACCTT GACAGGACCC
101 ACTACCGGCA AAGTGAACTC CTCCGTGAGC TGACACGTGGA CACCAAGATT
TGATGGCCGT TTCACATTGAG GAGGCACTCG ACGTGCACCT GTGGTCTAAA
151 CCTGCTACAG CGCCAGGAGC TCTTCCGTTG GGCCCCTGAG TCCTGTACAA
GGACGATGTC GCGGTCTCTG AGAAGGCAAC CGGGCAGTC AGGACATGTT
201 AAAGACTGTG TTCGTAGAGT TCACGGATCA ACTTTTCAGC GTTGCCAGGC
TTTCTGACAC AAGCATCTCA AGTGCCTAGT TGAAAAGTCG CAACGGTCCG
251 CCAGGCCACC ATGGATGGGT CTGCTGGGTG CTACCATCCA GGCTGAGGTT
GGTCCGGTGG TACCTACCCA GACGACCCAG GATGGTAGGT CCGACTCCAA
301 TACGACACGG TGGTCGTTAC CCTGAAGAAC ATGGCTTCTC ATCCCGTTAG
ATGCTGTGCC ACCAGCAATG GGACTTCTTG TACCGAAGAG TAGGGCAATC
351 TCTTCACGCT GTCGGCGTCT CCTTCTGGAA ATCTTCCGAA GGCCTGAAAT
AGAAAGTGCAG CAGCCGCAGA GGAAGACCTT TAGAAGGCTT CCGCGACTTA
401 ATGAGGATCA CACCAGCCAA AGGGAGAAGG AAGACGATAA AGTCCTTCCC
TACTCCTAGT GTGGTCGGTT TCCCTCTTC TTCTGCTATT TCAGGAAGGG
451 GGTAAAAGCC AACCTACGT CTGGCAGGTG CTGAAAGAAA ATGGTCCAAC
CCATTTCGG TTTGGATGCA CACCGTCCAG GACTTTCTT TACCAAGGTTG
501 AGCCTCTGAC CCACCATGTC TTACCTACTC ATACCTGTCT CACGTGGACC
TCGGAGACTG GGTGGTACAG AATGGATGAG TATGGACAGA GTGCACCTGG
551 TGGTGAAGA CCTGAATTG GGCCTCATG GAGCCCTGCT GTTTGTAGA
ACCACTTTCT GGACTTAAGC CCGGAGTAAC CTCGGGACGA CCAAACATCT
601 GAAGGGAGTC TGACCCAGAGA AAGGACCCAG AACCTGCACG AATTGTACT
CTTCCCTCAG ACTGGTCTCT TTCCCTGGTC TTGGACGTGC TAAACATGA
651 ACTTTTGCT GTCTTGATG AAGGGAAAAG TTGGCACTCA GCAAGAAATG
TGAAAAACGA CAGAAACTAC TTCCCTTTTC AACCGTGAGT CGTTCTTAC
701 ACTCCTGGAC ACGGGCCATG GATCCCGCAC CTGCCAGGGC CCAGCCTGCA
TGAGGACCTG TGCCCGGTAC CTAGGGCGTG GACGGTCCCG GTCGGACGT
751 ATGCACACAG TCAATGGCTA TGTCAACAGG TCTCTGCCAG GTCTGATCGG
TACGTGTGTC AGTTACCGAT ACAGTTGTCC AGAGACGGTC CAGACTAGCC
801 ATGTCATAAG AAATCAGTCT ACTGGCACGT GATTGGAATG GGCACCAAGCC
TACAGTATTG TTTAGTCAGA TGACCGTGCA CTAACCTTAC CCGTGGTCGG
851 CGGAAGTGCAG CTCCATTGTT CTTGAAGGCC ACACGTTCT CGTGAGGCAC
GCCTTCACGT GAGGTAAAAA GAACTCCGG TGTGCAAAGA GCAACTCCGTG
901 CATCGCCAGG CTTCCATTGGA GATCTGCCA CTAACCTTCC TCACTGCTCA
GTAGCGGTCC GAAGGAACCT CTAGAGCGGT GATTGAAAGG AGTGACGAGT
951 GACATTCTG ATGGACCTG GCCAGTTCT ACTGTTTGT CATATCTCTT
CTGTAAGGAC TACCTGGAAC CGGTCAAGGA TGACAAAACA GTATAGAGAA
1001 CCCACCACCA TGGTGGCATG GAGGCTCACG TCAGAGTAGA AAGCTGCGCC
GGGTGGTGGT ACCACCGTAC CTCCGAGTGC AGTCTCATCT TTGACGCGG
1051 GAGGAGCCCG AGCTCGGGAG GAAAGCTGAT GAAGAGGAAG ATTATGATGA
CTCCTCGGGG TCGACGCCCTC CTTTCGACTA CTTCTCCTTC TAATACTACT
1101 CAATTGTAC GACTCGGACA TGGACGTGGT CCGGCTCGAT GGTGACGACG
GTTAAACATG CTGAGCCTGT ACCTGCACCA GGCGAGCTA CCACGTGTC
1151 TGTCTCCCTT TATCCAAATC CGCTCGGTTG CCAAGAAGCA TCCCAAAACC
ACAGAGGGAA ATAGGTTAG GCGAGCCAAC GGTTCTTCGT AGGGTTTGG
1201 TGGGTGCACT ACATCTCTGC AGAGGAGGAG GACTGGGACT ACGCCCCCGC
ACCCACGTGA TGTAGAGACG TCTCCTCCTC CTGACCCCTGA TGCGGGGGCG
1251 GGTCCCCAGC CCCAGTGACA GAAGTTATAA AAGTCTCTAC TTGAACAGTG
CCAGGGGTGCG GGGTCACTGT CGGCAATATT TTCAGAGATG AACTTGTAC
1301 GTCCTCAGCG AATTGGTAGG AAATACAAAA AAGCTCGATT CGTCGCTTAC
CAGGAGTCGC TTAACCATCC TTTATGTTT TTGAGCTAA GCAGCGAATG

FIG. 5A

11/33

1351	ACGGATGTAA	CATTAAAGAC	TCGTAAAGCT	ATTCCGTATG	AATCAGGAAT
	TGCCCTACATT	GTAAATTCTG	AGCATTTCGA	TAAGGCATAC	TTAGTCCTTA
1401	CCTGGGACCT	TTACTTTATG	GAGAAGTTGG	AGACACACTT	TTGATTATAT
	GGACCCCTGGA	AATGAAATAC	CTCTTCAACC	TCTGTGTGAA	AACTAATATA
1451	TTAAGAATAA	AGCGAGCCGA	CCATATAACA	TCTACCCTCA	TGGAATCACT
	AATTCTTATT	TCGCTCGGCT	GGTATATTGT	AGATGGGAGT	ACCTTAGTGA
1501	GATGTCAGCG	CTTGCACCC	AGGGAGACTT	CTAAAAGGTT	GGAAACATT
	CTACAGTCGC	GAAACGTGGG	TCCCTCTGAA	GATTTTCAA	CCTTGTAAA
1551	GAAAGACATG	CCAATTCTGC	CAGGAGAGAC	TTTCAAGTAT	AAATGGACAG
	CTTTCTGTAC	GGTTAAGACG	GTCCTCTCTG	AAAGTTCAT	TTTACCTGTC
1601	TGACTGTGGA	AGATGGGCCA	ACCAAGTCCG	ATCCTCGGTG	CCTGACCCGC
	ACTGACACCT	TCTACCCGGT	TGGTTCAGGC	TAGGAGCCAC	GGACTGGCG
1651	TACTACTCGA	GCTCCATTAA	TCTAGAGAAA	GATCTGGCTT	CGGGACTCAT
	ATGATGAGCT	CGAGGTAAATT	AGATCTCTT	CTAGACCGAA	GCCCTGAGTA
1701	TGGCCCTCTC	CTCATCTGCT	ACAAAGAAC	TGTAGACCAA	AGAGGAAACC
	ACCGGGAGAG	GAGTAGACGA	TGTTTCTTAG	ACATCTGGTT	TCTCCTTGG
1751	AGATGATGTC	AGACAAGAGA	AACGTATCC	TGTTTCTGT	ATTGATGAG
	TCTACTACAG	TCTGTTCTCT	TTGCAGTAGG	ACAAAAGACA	TAAGCTACTC
1801	AATCAAAGCT	GGTACCTCGC	AGAGAATATT	CAGCGCTTCC	TCCCCAATCC
	TTAGTTTCGA	CCATGGAGCG	TCTCTTATAA	GTCGCGAAGG	AGGGGTTAGG
1851	GGATGGATTA	CAGCCCCAGG	ATCCAGAGTT	CCAAGCTTCT	AACATCATGC
	CCTACCTAAT	GTCGGGGTCC	TAGGTCTCAA	GGTCGAAGA	TTGTAGTACG
1901	ACAGCATCAA	TGGCTATGTT	TTTGATAGCT	TGCACTGTC	GGTTGTTTG
	TGTCGTAGTT	ACCGATACAA	AAACTATCGA	ACGTCGACAG	CCAAACAAAC
1951	CACGAGGTGG	CATACTGGTA	CATTCTAAGT	GTTGGAGCAC	AGACGGACTT
	GTGCTCCACC	GTATGACCAT	GTAAGATTCA	CAACCTCGTG	TCTGCCTGAA
2001	CCTCTCCGTC	TTCTTCTCTG	GCTACACCTT	CAAACACAAA	ATGGTCTATG
	GGAGAGGCAG	AAGAAGAGAC	CGATGTGGAA	GTTTGTGTT	TACCAGATAC
2051	AAGACACACT	CACCCTGTTC	CCCTTCTCAG	GAGAAACGGT	CTTCATGTCA
	TTCTGTGTGA	GTGGGACAAG	GGGAAGAGTC	CTCTTGCCA	GAAGTACAGT
2101	ATGGAAAACC	CAGGTCTCTG	GGTCCTTGGG	TGCCACAAC	CAGACTTGCG
	TACCTTTGG	GTCCAGAGAC	CCAGGAACCC	ACGGTGTGA	GTCTGAACGC
2151	GAACAGAGGG	ATGACAGCCT	TACTGAAGGT	GTATAGTTGT	GACAGGGACA
	CTTGTCTCCC	TACTGTCGGA	ATGACTTCCA	CATATCAACA	CTGTCCCTGT
2201	TTGGTGATTA	TTATGACAAC	ACTTATGAAG	ATATTCCAGG	CTTCTTGCTG
	AACCACTAAT	AATACTGTG	TGAATACTTC	TATAAGGTCC	GAAGAACGAC
2251	AGTGGAAAGA	ATGTCATTGA	ACCTAGGAGC	TTTGCCCAGA	ATTCAAGACC
	TCACCTTTCT	TACAGTAAC	TGGATCCTCG	AAACGGGTCT	TAAGTTCTGG
2301	CCCTAGTGGG	AGCGCTCCAA	AGCCTCCGGT	CCTGCGACGG	CATCAGAGGG
	GGGATCACGC	TCGCGAGGGT	TCGGAGGCCA	GGACGCTGCC	GTAGTCTCCC
2351	ACATAAGCCT	TCCTACTTT	CAGCCGGAGG	AAGACAAAAT	GGACTATGAT
	TGTATTCGGA	AGGATGAAA	GTCGGCCTCC	TTCTGTTTA	CCTGATACTA
2401	GATATCTCT	CAACTGAAAC	GAAGGGAGAA	GATTTTGACA	TTTACGGTGA
	CTATAGAAGA	GTTGACTTTG	CTTCCCTCTT	CTAAAACGT	AAATGCCACT
2451	GGATGAAAAT	CAGGACCCCTC	GCAGCTTCA	GAAGAGAAC	CGACACTATT
	CCTACTTTA	GTCCTGGGAG	CGTCGAAAGT	CTTCTCTTGG	GCTGTGATAA
2501	TCATTGCTGC	GGTGGAGCAG	CTCTGGGATT	ACGGGATGAG	CGAATCCCCC
	AGTAACGACG	CCACCTCGTC	GAGACCCCAA	TGCCCTACTC	GCTTAGGGGG
2551	CGGGCGCTAA	GAAACAGGGC	TCAGAACCGA	GAGGTGCCTC	GGTCAGAGAA
	GCCCGCGATT	CTTGTCCTCG	AGTCTTGCCT	CTCCACGGAG	CCAAGTTCTT
2601	GGTGGTCTTC	CGGGAATTG	CTGACGGCTC	CTTCACGCGAG	CCGTCGTACC
	CCACCCAGAAG	GCCCTTAAAC	GACTGCCGAG	GAAGTGCCTC	GGCAGCATGG
2651	GCGGGGAAC	CAACAAACAC	TTGGGGCTCT	TGGGACCCCTA	CATCAGAGCG
	CGCCCCCTGA	GTTGTTTGTG	AACCCCGAGA	ACCCTGGGAT	GTAGTCTCGC
2701	GAAAGTTGAAG	ACAACATCAT	GGTAACCTTC	AAAAACCAGG	CGTCTCGTCC
	CTTCAACTTC	TGTTGTAGTA	CCATTGAAAG	TTTTGGTCC	GCAGAGCAGG
2751	CTATTCTTC	TACTCGAGCC	TTATTCTTA	TCCGGATGAT	CAGGAGCAAG

FIG. 5B

12/33

2801	GATAAGGAAG	ATGAGCTCGG	AATAAAGAAT	AGGCCTACTA	GTCCTCGTTC
	GGGCAGAAC	TCGACACAAC	TTCGTCCAGC	CAAATGAAAC	CAGAACTTAC
2851	CCCCTCTGG	AGCTGTGTT	AAGCAGGTCG	GTTTACTTTG	GTCTTGAATG
	TTTGAAAG	TGCAGCATCA	CATGGCACCC	ACAGAAGACG	AGTTTGACTG
2901	AAAACCTTC	ACGTCGTAGT	GTACCGTGGG	TGTCTTCTGC	TCAAACGTAC
	CAAAGCTGG	GCCTACTTT	CTGATGTTGA	CCTGGAAAAA	GATGTGCACT
2951	GTTCGGGACC	CGGATGAAA	GACTACAAC	GGACCTTTT	CTACACGTGA
	CAGGCTTGT	CGGCCCCCTT	CTGATCTGCC	GCGCCAACAC	CCTGAACGCT
3001	GTCCGAACTA	GCCGGGGAA	GACTAGACGG	CGCGGTTGTG	GGACTTGCGA
	GCTCACGGTA	GACAAGTGAC	CGTGCAAGAA	TTTGCTCTGT	TTTCACTAT
3051	CGAGTCCAT	CTGTTCACTG	GCACGTTCTT	AAACGAGACA	AAAAGTGATA
	TTTGATGAG	ACAAAGAGCT	GGTACTTCAC	TGAAAATGTG	GAAAGGAACT
3101	AAAACACTC	TGTTTCTCGA	CCATGAAGTG	ACTTTACAC	CTTCCTTGA
	GCCGGGCC	CTGCCATCTG	CAGATGGAGG	ACCCCAC	GAAAGAAAAC
3151	CGGCCCGGG	GACGGTAGAC	GTCTACCTCC	TGGGGTGAGA	CTTCTTTG
	TATCGCTTC	ATGCAATCAA	TGGCTATGTG	ATGGATACAC	TCCCTGGCTT
	ATAGCGAAGG	TACGTTAGTT	ACCGATACAC	TACCTATGTG	AGGGACCGAA
3201	AGTAATGGCT	CAGAATCAA	GGATCCGATG	GTATCTGCTC	AGCATGGGCA
	TCATTACCGA	GTCTTAGTT	CCTAGGCTAC	CATAGACGAG	TCGTACCCGT
3251	GCAATGAAA	TATCCATTG	ATTCAATTAA	GC GGACACGT	GTTCACTGTG
	CGTTACTTT	ATAGGTAAGC	TAAGTAAAAT	CGCCTGTGCA	CAAGTCACAT
3301	CGGAAAAAGG	AGGAGTATAA	AATGGCCGTG	TACAATCTCT	ATCCGGGTGT
	GCCTTTTCC	TCCTCATATT	TTACCGGCAC	ATGTTAGAGA	TAGGCCACAA
3351	CTTGAGACA	GTGGAAATGC	TACCGTCCAA	AGTTGGAATT	TGGCGAATAG
	GAAACTCTGT	CACCTTACG	ATGGCAGGTT	TCAACCTTAA	ACCGCTTATC
3401	AATGCCTGAT	TGGCGAGCAC	CTGCAAGCTG	GGATGAGCAC	GACTTCCTG
	TTACGGACTA	ACCGCTCGT	GACGTTCGAC	CCTACTCGT	CTGAAAGGAC
3451	GTGTACAGCA	AGAAGTGTCA	GACTCCCCTG	GGAATGGCTT	CTGGACACAT
	CACATGTCGT	TCTTCACAGT	CTGAGGGGAC	CCTTACCGAA	GACCTGTGTA
3501	TAGAGATT	CAGATTACAG	CTTCAGGACA	ATATGGACAG	TGGGCCCAA
	ATCTCTAAA	GTCTAATGTC	GAAGTCCTGT	TATACCTGTC	ACCCGGGGTT
3551	AGCTGGCCAG	ACTTCATTAT	TCCGGATCAA	TCAATGCCTG	GAGCACCAAG
	TCGACCGGTC	TGAAGTAATA	AGGCCTAGTT	AGTTACGGAC	CTCGTGGTTC
3601	GAGCCCTTT	CTTGGATCAA	GGTGGATCTG	TTGGCACCAA	TGATTATTCA
	CTCGGGAAA	GAACCTAGTT	CCACCTAGAC	AACCGTGGTT	ACTAATAAGT
3651	CGGCATCAAG	ACCCAGGGTG	CCCCTCAGAA	GTTCTCCAGC	CTCTACATCT
	GCCGTAGTTC	TGGGTCCCAC	GGGCAGTCTT	CAAGAGGTG	GAGATGTAGA
3701	CTCAGTTAT	CATCATGTAT	AGTCTTGATG	GGAAGAAGTG	GCAGACTTAT
	GAGTCAAATA	GTAGTACATA	TCAGAACTAC	CCTTCTTCAC	CGTCTGAATA
3751	CGAGGAAATT	CCACTTGAAC	CTTAATGGTC	TTCTTGGCA	ATGTGGATT
	GCTCCTTAA	GGTGACCTG	GAATTACAG	AAGAAACCGT	TACACCTAAG
3801	ATCTGGGATA	AAACACAATA	TTTTAACCC	TCCAATTATT	GCTCGATACA
	TAGACCTAT	TTTGTGTTAT	AAAAATTGGG	AGGTTAATAA	CGAGCTATGT
3851	TCCGTTGCA	CCCAACTCAT	TATAGCATTC	GCAGCACTCT	TCGCATGGAG
	AGGCAAACGT	GGGTTGAGTA	ATATCGTAAG	CGTCGTCA	AGCGTACCTC
3901	TTGATGGGCT	GTGATTTAAA	TAGTTGCAGC	ATGCCATTGG	GAATGGAGAG
	AACTACCCGA	CACTAAATT	ATCAACGTG	TACGGTAACC	CTTACCTCTC
3951	TAAAGCAATA	TCAGATGCAC	AGATTACTGC	TTCATCCTAC	TTTACCAATA
	ATTTCGTTAT	AGTCTACGTG	TCTAATGACG	AAGTAGGATG	AAATGGTTAT
4001	TGTTTGCAC	CTGGTCTCCT	TCAAAAGCTC	GACTTCACCT	CCAAGGGAGG
	ACAAACGGTG	GACCAGAGGA	AGTTTTCGAG	CTGAAGTGG	GGTCCCTCC
4051	AGTAATGCT	GGAGACCTCA	GGTGAATAAT	CCAAAAGAGT	GGCTGCAAGT
	TCATTACCGA	CCTCTGGAGT	CCACTTATT	GGTTTCTCA	CCGACGTTCA
4101	GGACTTCCAG	AAGACAATGA	AAGTCACAGG	AGTAACTACT	CAGGGAGTAA
	CCTGAAGGTC	TTCTGTTACT	TTCACTGTCC	TCATTGATGA	GTCCCTCATT
4151	AATCTCTGCT	TACCAGCATG	TATGTGAAGG	AGTTCCCTCAT	CTCCAGCAGT
	TTAGAGACGA	ATGGTCGTAC	ATACACTTCC	TCAAGGAGTA	GAGGTCGTCA

FIG. 5C

13/33

4201 CAAGATGGCC ATCAGTGGAC TCTCTTTTT CAGAATGGCA AAGTAAAGGT
GTTCTACCGG TAGTCACCTG AGAGAAAAAA GTCTTACCGT TTCATTTCCA
4251 TTTTCAGGGA AATCAAGACT CCTTCACACC TGTGGTGAAC TCTCTAGACC
AAAAGTCCCT TTAGTTCTGA GGAAGTGTGG ACACCACTTG AGAGATCTGG
4301 CACCGTTACT GACTCGCTAC CTTCGAATTG ACCCCCAGAG TTGGGTGCAC
GTGGCAATGA CTGAGCGATG GAAGCTTAAG TGGGGGTCTC AACCCACGTG
4351 CAGATTGCC TGAGGATGGA GGTCTGGGC TGCGAGGCAC AGGACCTCTA
GTCTAACGGG ACTCCTACCT CCAAGACCCG ACGCTCCGTG TCCTGGAGAT
4401 C
G

1-57 SIGNAL PEPTIDE
58-1173 A1 DOMAIN
1174-2277 A2 DOMAIN
2278-2349 OL LINKER
2350-3462 ap-A3 DOMAINS
3463-3921 C1 DOMAIN
3922-4401 C2 DOMAIN

FIG. 5D

AMINO ACID SEQUENCE OF HP46/SQ

1	MQLELSTCVF	LCLLPLGFSA	IRYYLGAIVE	LSWDYRQSEL	LRELHVDTRF
51	PATAPGALPL	GPSVLYKKTV	FVEFTDQLFS	VARPRPPWMG	LLGPTIQAEV
101	YDTVVVTLKN	MASHPVSLHA	VGVSFWKSSE	GAEYEDHTSQ	REKEDDKVLP
151	GKSQTYVWQV	LKENGPTASD	PPCLTYSYLS	HVDLVKDLNS	GLIGALLVCR
201	EGSLTRERTQ	NLHEFVLLFA	VFDEGKSWHS	ARNDSWTRAM	DPAPARAQPA
251	MHTVNGYVNR	SLPGLIGCHK	KSVYWHVIGM	GTSPEVHSIF	LEGHTFLVRH
301	HRQASLEISP	LTFLTAQTFL	MDLGQFLLLFC	HISSHGGM	EAHVRVESCA
351	EEPQLRRKAD	EEEDYDDNLY	DSDMDVVRLD	GDDVSPFIQI	RSVAKKHPKT
401	WVHYIAAEEE	DWDYAPLVLA	PDDRSYKSQY	LNNGPQRIGR	KYKKVRFMAY
451	TDETFKTREA	IQHESGILGP	LLYGEVGDTL	LIIFKNQASR	PYNIYPHGIT
501	DVRPLYSRRL	PKGVKHLKDF	PILPGEIFKY	KWTVTVEDGP	TKSDPRCLTR
551	YYSSFVNMER	DLASGLIGPL	LICYKESVDO	RGNQIMSDKR	NVILFSVFDE
601	NRSWYLTENI	QRFLPNPAGV	QLEDPEFQAS	NIMHSINGYV	FDSLQLSVCL
651	HEVAYWYIILS	IGAQTDLFSV	FFSGYTFKHK	MVYEDTLTLF	PFSGETVFMS
701	MENPGLWILG	CHNSFLRNRG	MTALLKVSSC	DKNTGDYYED	SYEDISAYLL
751	SKNNAIEPRS	FSQNPPVLKR	HQREITRTTL	QSDQEEIDYD	DTISVEMKKE
801	DFDIYDEDEN	QSPRSFQKKT	RHYFIAAVER	LWDYGMSSSP	HVLRNRAQSG
851	SVPQFKVVF	QEFTDGSFTQ	PLYRGELNEH	LGLLGPYIRA	EVEDNIMVTF
901	RNQASRPYSF	YSSLISYEED	QRQGAEPRKN	FVKPNETKTY	FWKVQHHMAP
951	TKDEFDCKAW	AYFSDVDLEK	DVHSGLIGPL	LVCHTNLNP	AHGRQVTVOE
1001	FALFFTIFDE	TKSWYFTENM	ERNCRAPCNI	QMEDPTFKEN	YRFHAINGYI
1051	MDTPGLVMA	QDQRIRWYLL	SMGSNENIHS	IHFSGHVFTV	RKKEEYKMAL
1101	YNLYPGVFET	VEMLPSKAGI	WRVECLIGEH	LHAGMSTLFL	VYSNKCQTPL
1151	GMASGHIRDF	QITASGQYQGQ	WAPKLARLHY	SGSINAWSTK	EPFSWIKVDL
1201	LAPMIIHGIK	TQGARQKFSS	LYISQFIIMY	SLDGKKWQTY	RGNSTGTLMV
1251	FFGNVDSSGI	KHNI FNPPII	ARYIRLHPTH	YSIRSTLRLME	LMGCDLNSCS
1301	MPLGMESKAI	SDAQITASSY	FTNMFATWSP	SKARLHLOGR	SNAWRPQVNN
1351	PKEWLQVDFQ	KTMKVTGVTT	QGVKSLLTSM	YVKEFLISSS	QDGHQWTLF
1401	QNGKVKVFQG	NQDSFTPVVN	SLDPPLLTRY	LRIHPQSWVH	QIALRMEVLG
1451	CEAQDLY*				

1-19 SIGNAL PEPTIDE
 20-391 A1 DOMAIN
 392-759 A2 DOMAIN
 760-773 SQ LINKER
 774-1144 ap-A3
 1145-1297 C1 DOMAIN
 1298-1457 C2 DOMAIN

FIG. 6

HP46 / SQ NUCLEOTIDE SEQUENCE

1	ATGCAGCTAG	AGCTCTCCAC	CTGTGTCTT	CTGTGTCTCT	TGCCACTCGG
	TACGTCGATC	TCGAGAGGTG	GACACAGAAA	GACACAGAGA	ACGGTGAGCC
51	CTTAGTGCC	ATCAGGAGAT	ACTACCTGGG	CGCAGTGGAA	CTGTCCTGGG
	GAAATCACGG	TAGTCCTCTA	TGATGGACCC	GCGTCACCTT	GACAGGACCC
101	ACTACCGGCA	AAGTGAACTC	CTCCGTGAGC	TGCACGTGGA	CACCAAGATT
	TGATGGCCGT	TTCACTTGAG	GAGGCACTCG	ACGTGCACCT	GTGGTCTAAA
151	CCTGCTACAG	CGCCAGGAGC	TCTTCCGTTG	GGCCCGTCAG	TCCTGTACAA
	GGACGATGTC	GCGGTCTCG	AGAAGGCAAC	CCGGGCAGTC	AGGACATGTT
201	AAAGACTGTG	TTCGTAGAGT	TCACGGATCA	ACTTTCAGC	GTTGCCAGGC
	TTTCTGACAC	AAGCATCTCA	AGTGCCTAGT	TGAAAAGTCG	CAACGGTCCG
251	CCAGGCCACC	ATGGATGGGT	CTGCTGGGTC	CTACCATCCA	GGCTGAGGTT
	GGTCCGGTGG	TACCTACCCA	GACGACCCAG	GATGGTAGGT	CCGACTCCAA
301	TACGACACGG	TGGTCGTTAC	CCTGAAGAAC	ATGGCTTCTC	ATCCCCTTAG
	ATGCTGTGCC	ACCAGCAATG	GGACTTCTTG	TACCGAAGAG	TAGGGCAATC
351	TCTTCACGCT	GTCGGCGTCT	CCTTCTGGAA	ATCTTCCGAA	GGCGCTGAAT
	AGAAGTGCAGA	CAGCCGCAGA	GGAAGACCTT	TAGAAGGCTT	CCGCGACTTA
401	ATGAGGATCA	CACCAGCCAA	AGGGAGAAGG	AAGACGATAA	AGTCCTTCCC
	TACTCCTAGT	GTGGTCGGTT	TCCCTCTTC	TTCTGCTATT	TCAGGAAGGG
451	GGTAAAAGCC	AAACCTACGT	CTGGCAGGTC	CTGAAAGAAA	ATGGTCCAAC
	CCATTTTCGG	TTTGGATGCA	GACCGTCCAG	GACTTTCTT	TACCAAGGTTG
501	AGCCTCTGAC	CCACCATGTC	TTACCTACTC	ATACCTGTCT	CACGTGGACC
	TCGGAGACTG	GGTGGTACAG	AATGGATGAG	TATGGACAGA	GTGCACCTGG
551	TGGTGAAGAAGA	CCTGAATTG	GGCCTCATG	GAGCCCTGCT	GGTTGTAGA
	ACCACTTTCT	GGACTTAAGC	CCGGAGTAAC	CTCGGGACGA	CCAAACATCT
601	GAAGGGAGTC	TGACCAGAGA	AAGGACCCAG	AACCTGCACG	AATTGTACT
	CTTCCCTCAG	ACTGGTCTCT	TTCCTGGGTC	TTGGACGTGC	TTAACACATGA
651	ACTTTTGCT	GTCTTTGATG	AAGGGAAAAG	TTGGCACTCA	GCAAGAAATG
	TGAAAAAACGA	CAGAAACTAC	TTCCCTTTTC	AACCGTGAGT	CGTTCTTAC
701	ACTCCTGGAC	ACGGGCCATG	GATCCCGCAC	CTGCCAGGGC	CCAGCCTGCA
	TGAGGACCTG	TGCCCGGTAC	CTAGGGCGTG	GACGGTCCCC	GGTCGGACGT
751	ATGCACACAG	TCAATGGCTA	TGTCAACAGG	TCTCTGCCAG	GTCTGATCGG
	TACGTGTGTC	AGTTACCGAT	ACAGTTGTCC	AGAGACGGTC	CAGACTAGCC
801	ATGTCATAAG	AAATCAGTCT	ACTGGCACGT	GATTGGAATG	GGCACCAAGCC
	TACAGTATTC	TTTAGTCAGA	TGACCGTGCA	CTAACCTTAC	CCGTGGTCGG
851	CGGAAGTGCA	CTCCATTTT	CTTGAAGGCC	ACACGTTCT	CGTGAGGCAC
	GCCTTCACGT	GAGGTAAAAA	GAACCTCCGG	TGTGCAAAGA	GCACTCCGTG
901	CATGCCAGG	CTTCCTTGG	GATCTCGCCA	CTAACTTTCC	TCACTGCTCA
	GTAGCGGTCC	GAAGGAACCT	CTAGAGCGGT	GATTGAAAGG	AGTGACGAGT
951	GACATTCTG	ATGGACCTTG	GCCAGTTCT	ACTGTTTGT	CATATCTCTT
	CTGTAAGGAC	TACCTGGAAC	CGGTCAAGGA	TGACAAAACA	GTATAGAGAA
1001	CCCACCACCA	TGGTGGCATG	GAGGCTCACG	TCAAGAGTAGA	AAGCTGCGCC
	GGGTGGTGGT	ACCACCGTAC	CTCCGAGTGC	AGTCTCATCT	TTCGACGCCG
1051	GAGGAGCCCC	AGCTGCGGAG	GAAAGCTGAT	GAAGAGGAAG	ATTATGATGA
	CTCCTCGGGG	TCGACGCCTC	CTTCGACTA	CTTCTCCTTC	TAATACTACT
1101	CAATTGTAC	GACTCGGACA	TGGACGTGGT	CCGGCTCGAT	GGTGACGACG
	GTAAACATG	CTGAGCCTGT	ACCTGCACCA	GGCCGAGCTA	CCACTGCTGC
1151	TGTCTCCCTT	TATCCAATC	CGCTCAGTTG	CCAAGAAGCA	TCCTAAAAC
	ACAGAGGGAA	ATAGGTTAG	GCGAGTCAAC	GGTTCTTCGT	AGGATTTGA
1201	TGGGTACATT	ACATTGCTGC	TGAAGAGGAG	GACTGGGACT	ATGCTCCCTT

FIG. 7A

1251	ACCCATGTAA	TGTAACGACG	ACTTCTCCTC	CTGACCCTGA	TACGAGGGAA
	AGTCCTCGCC	CCCGATGACA	GAAGTTATAA	AAGTCAATAT	TTGAACAAATG
1301	TCAGGAGCGG	GGGCTACTGT	CTTCAATATT	TTCAGTTATA	AACTTGTAC
	GCCCTCAGCG	GATTGGTAGG	AAGTACAAAA	AAGTCCGATT	TATGGCATAAC
1351	CGGGAGTCGC	CTAACCATCC	TTCATGTTT	TTCAGGCTAA	ATACCGTATG
	ACAGATGAAA	CCTTTAAGAC	GCCTGAAGCT	ATTCAAGCATG	AATCAGGAAT
1401	TGTCTACTTT	GGAAATTCTG	CGCACTTCGA	TAAGTCGTAC	TTAGTCCTTA
	CTTGGGACCT	TTACTTTATG	GGGAAGTTGG	AGACACACTG	TTGATTATAT
1451	GAACCTGGA	AATGAAATAC	CCCTTCAACC	TCTGTGTGAC	AACTAATATA
	TTAAGAATCA	AGCAAGCAGA	CCATATAACA	TCTACCCCTCA	CGGAATCACT
1501	AATTCTTAGT	TCGTTCGTCT	GGTATATTGT	AGATGGGAGT	GCCTTAGTGA
	GATGTCCGTC	CTTTGTATTTC	AAGGAGATTA	CCAAAAGGTG	TAAAACATT
1551	CTACAGGCAG	GAAACATAAG	TTCCTCTAAT	GGTTTTCCAC	ATTTGTAAA
	GAAGGATTT	CCAATTCTGC	CAGGAGAAAT	ATTCAAATAT	AAATGGACAG
1601	CTTCCTAAAA	GGTTAAGACG	GTCCTCTT	TAAGTTTATA	TTTACCTGTC
	TGACTGTAGA	AGATGGGCCA	ACTAAATCAG	ATCCCGGGTG	CCTGACCCGC
1651	ACTGACATCT	TCTACCCGGT	TGATTTAGTC	TAGGCGCCAC	GGACTGGGCG
	TATTACTCTA	GTTCGTTAA	TATGGAGAGA	GATCTAGCTT	CAGGACTCAT
1701	ATAATGAGAT	CAAAGCAATT	ATACCTCTCT	CTAGATCGAA	GTCCTGAGTA
	TGGCCCTCTC	CTCATCTGCT	ACAAAGAAC	TGTAGTCCAA	AGAGGAAACC
1751	ACCGGGAGAG	GAGTAGACGA	TGTTTCTTAG	ACATCTAGTT	TCTCCTTGG
	AGATAATGTC	AGACAAGAGG	AATGTCATCC	TGTTTCTGT	ATTGATGAG
1801	TCTATTACAG	TCTGTTCTCC	TTACAGTAGG	ACAAAAGACA	TAAACTACTC
	AACCGAAGCT	GGTACCTCAC	AGAGAATATA	CAACGCTTTC	TCCCCAATCC
1851	TTGGCTTCGA	CCATGGAGTG	TCTCTTATAT	GTTGCGAAAG	AGGGGTTAGG
	AGCTGGAGTA	CAGCTTGAGG	ATCCAGAGTT	CCAAGCCTCC	AACATCATGC
1901	TCGACCTCAC	GTCGAACCTC	TAGGTCTCAA	GGTCGAAGG	TTGTAGTACG
	ACAGCATCAA	TGGCTATGTT	TTTGATAGTT	TGCAGTTGTC	AGTTGTTTG
1951	TGTCGTAGTT	ACCGATACAA	AAACTATCAA	ACGTCAACAG	TCAAACAAAC
	CATGAGGTGG	CATACTGGTA	CATTCTAAC	ATTGGAGCAC	AGACTGACTT
2001	GTACTCCACC	GTATGACCAT	GTAAGATTG	TAACCTCGTG	TCTGACTGAA
	CCTTCTGTC	TTCTTCTCTG	GATATACCTT	CAAACACAAA	ATGGTCTATG
2051	GGAAAGACAG	AAGAAGAGAC	CTATATGGAA	GTTGTGTTT	TACCAAGATA
	AAGACACACT	CACCCATTTC	CCATTCTCAG	GAGAAACTGT	CTTCATGTCG
2101	TTCTGTGTGA	GTGGGATAAG	GGTAAGAGTC	CTCTTGACA	GAAGTACAGC
	ATGGAAAACC	CAGGTCTATG	GATTCTGGGG	TGCCACAAC	CAGACTTCG
2151	TACCTTTGG	GTCCAGATAC	CTAACGACCC	ACGGTGTGTA	GTCTGAAAGC
	GAACAGAGGC	ATGACCGCCT	TACTGAAGGT	TTCTAGTTGT	GACAAGAAC
2201	CTTGTCTCCG	TACTGGCGGA	ATGACTTCCA	AAGATCAACA	CTGTTCTTGT
	CTGGTGATTA	TTACGAGGAC	AGTTATGAAG	ATATTTCAGC	ATACTTGCTG
2251	GACCACTAAT	AATGCTCCTG	TCAATACTTC	TATAAAGTCG	TATGAACGAC
	AGTAAAAACA	ATGCCATTGA	ACCTAGGAGC	TTCTCTCAGA	ATCCACCAAGT
2301	TCATTTTGTT	TACGGTAACT	TGGATCCTCG	AAGAGAGTCT	TAGGTGGTCA
	CTTGAAACGC	CATCAACGGG	AAATAACTCG	TACTACTCTT	CAGTCAGATC
2351	GAACATTGCG	GTAGTTGCC	TTTATTGAGC	ATGATGAGAA	GTCAGTCTAG
	AAGAGGAAAT	TGACTATGAT	GATACCATAT	CAGTTGAAAT	GAAGAAGGAA
2401	TTCTCCTTA	ACTGATACTA	CTATGGTATA	GTCAAACCTTA	CTTCTCCTT
	GATTTTGACA	TTTATGATGA	GGATGAAAAT	CAGAGCCCCC	GCAGCTTTCA
2451	CTAAAACGT	AAATACTACT	CCTACTTTA	GTCTCGGGGG	CGTCGAAAGT
	AAAGAAAACA	CGACACTATT	TTATTGCTGC	AGTGGAGAGG	CTCTGGGATT
2501	TTTCTTTGT	GCTGTGATAA	AATAACGACG	TCACCTCTCC	GAGACCCTAA
	ATGGGATGAG	TAGCTCCCCA	CATGTTCTAA	GAAACAGGGC	TCAGAGTGGC

17/33

2551	TACCCTACTC	ATCGAGGGGT	GTACAAGATT	CTTGTCCCG	AGTCTCAGCG
	AGTGTCCCTC	AGTTCAAGAA	AGTTGTTTC	CAGGAATTAA	CTGATGGCTC
2601	TCACAGGGAG	TCAAGTTCTT	TCAACAAAAG	GTCCTTAAAT	GACTACCGAG
	CTTTACTCAG	CCCTTATACC	GTGGAGAACT	AAATGAACAT	TTGGGACTCC
2651	GAAATGAGTC	GGGAATATGG	CACCTCTTGA	TTTACTTGTAA	AACCCTGAGG
	TGGGGCCATA	TATAAGAGCA	GAAGTTGAAG	ATAATATCAT	GGTAACTTTC
2701	ACCCCGGTAT	ATATTCTCGT	CTTCAACTTC	TATTATAGTA	CCATTGAAAG
	AGAAAATCAGG	CCTCTCGTCC	CTATTCTTC	TATTCTAGCC	TTATTTCTTA
2751	TCTTTAGTCC	GGAGAGCAGG	GATAAGGAAG	ATAAGATCGG	AATAAAAGAAT
	TGAGGAAGAT	CAGAGGCAAG	GAGCAGAAC	TAGAAAAAAAC	TTTGTCAAGC
2801	ACTCCTTCTA	GTCTCCGTT	CTCGTCTTGG	ATCTTTTTTG	AAACAGTTCG
	CTAATGAAAC	CAAAACTTAC	TTTGGAAAG	TGCAACATCA	TATGGCACCC
2851	GATTACTTTG	GTTTGAATG	AAAACCTTTC	ACGTTGTAGT	ATACCGTGGG
	ACTAAAGATG	AGTTGACTG	CAAAGCCTGG	GCTTATTCT	CTGATGTTGA
2901	TGATTTCTAC	TCAAACGTAC	GTTCCGGACC	CGAATAAAGA	GACTACAACT
	CCTGGAAAAA	GATGTGCACT	CAGGCCTGAT	TGGACCCCTT	CTGGTCTGCC
2951	GGACCTTTT	CTACACGTGA	GTCCGGACTA	ACCTGGGGAA	GACCAGACGG
	ACACTAACAC	ACTGAACCC	GCTCATGGGA	GACAAGTGC	AGTACAGGAA
3001	TGTGATTGTG	TGACTTGGGA	CGAGTACCC	CTGTTCACTG	TCATGTCATT
	TTTGCTCTGT	TTTCACCAT	CTTGATGAG	ACCAAAAGCT	GGTACTTCAC
3051	AAACGAGACA	AAAAGGGTA	GAAACTACTC	TGGTTTTCGA	CCATGAAGTG
	TGAAAATATG	GAAAGAAACT	GCAGGGCTCC	CTGCAATATC	CAGATGGAAG
3101	ACTTTTATAC	CTTCTTTGA	CGTCCCAGG	GACGTTATAG	GTCTACCTTC
	ATCCCAC	TAAAGAGAA	TATCGCTTCC	ATGCAATCAA	TGGCTACATA
3151	TAGGGTGAAA	ATTCTCTTA	ATAGCGAAGG	TACGTTAGTT	ACCGATGTAT
	ATGGATACAC	TACCTGGCTT	AGTAATGGCT	CAGGATCAA	GGATTGATG
3201	TACCTATGTG	ATGGACCGAA	TCATTACCGA	GTCCTAGTT	CCTAAGCTAC
	GTATCTGCTC	AGCATGGGCA	GCAATGAAAA	CATCCATTCT	ATTCAATTCA
3251	CATAGACGAG	TCGTACCCGT	CGTTACTTTT	GTAGGTAAGA	TAAGTAAAGT
	GTGGACATGT	GTTCACTGTA	CGAAAAAAAG	AGGAGTATAA	AATGGCACTG
3301	CACCTGTACA	CAAGTGACAT	GCTTTTTTC	TCCTCATATT	TTACCGTGCAC
	TACAATCTCT	ATCCAGGTGT	TTTGAGACA	GTGGAATGT	TACCATCCAA
3351	ATGTTAGAGA	TAGGTCCACA	AAAACTCTGT	CACCTTACA	ATGGTAGGTT
	AGCTGGAATT	TGGCGGGTGG	AATGCCTTAT	TGGCGAGCAT	CTACATGCTG
3401	TCGACCTTAA	ACCGCCCACC	TTACGGAATA	ACCGCTCGA	GATGTACGAC
	GGATGAGCAC	ACTTTTCTG	GTGTACAGCA	ATAAGTGTCA	GACTCCCCGT
3451	CCTACTCGTG	TGAAAAAGAC	CACATGTCGT	TATTACAGT	CTGAGGGGAC
	GGAATGGCTT	CTGGACACAT	TAGAGATTT	CAGATTACAG	CTTCAGGACA
3501	CCTTACCGAA	GACCTGTGTA	ATCTCTAAAA	GTCTAATGTC	GAAGTCCTGT
	ATATGGACAG	TGGGCCCAA	AGCTGGCCAG	ACTTCATTAT	TCCGGATCAA
3551	TATACCTGTC	ACCCGGGGTT	TCGACCGGT	TGAAGTAATA	AGGCCTAGTT
	TCAATGCCTG	GAGCACCAAG	GAGCCCTTT	CTTGGATCAA	GGTGGATCTG
3601	AGTTACGGAC	CTCGTGGTC	CTCGGGAAAA	GAACCTAGTT	CCACCTAGAC
	TTGGCACCAA	TGATTATTCA	CGGCATCAAG	ACCCAGGGTG	CCCGTCAGAA
3651	AACC GTGGTT	GCCGTAGTTC	GCCGTAGTTC	TGGGTCCCAC	GGGCAGTCTT
	GTTCTCCAGC	CTCTACATCT	CTCAGTTTAT	CATCATGTAT	AGTCTTGATG
3701	CAAGAGGTG	GAGATGTAGA	GAGTCAAATA	GTAGTACATA	TCAGAACTAC
	GGAAGAAGTG	GCAGACTTAT	CGAGGAAATT	CCACTGGAAC	CTTAATGGTC
3751	CCTTCTTCAC	CGTCTGAATA	GCTCCTTAA	GGTGACCTG	GAATTACCAAG
	TTCTTTGGCA	ATGTGGATTC	ATCTGGGATA	AAACACAATA	TTTTAACCC
3801	AAGAAACCGT	TACACCTAAG	TAGACCCTAT	TTTGTGTTAT	AAAAATTGGG
	TCCAATTATT	GCTCGATACA	TCCGTTTGCA	CCCAACTCAT	TATAGCATTG

FIG. 7C

18/33

3851	AGGTTAATAA	CGAGCTATGT	AGGCAAAGCT	GGGTTGAGTA	ATATCGTAAG
	GCAGCACTCT	TCGCATGGAG	TTGATGGGCT	GTGATTAAA	TAGTTGCAGC
	CGTCGTGAGA	AGCGTACCTC	AACTACCCGA	CACTAAATT	ATCAACGTCG
3901	ATGCCATTGG	GAATGGAGAG	TAAAGCAATA	TCAGATGCAC	AGATTACTGC
	TACGGTAACC	CTTACCTCTC	ATTCGTTAT	AGTCTACGTG	TCTAATGACG
3951	TTCATCCTAC	TTTACCAATA	TGTTTGCCAC	CTGGTCTCCT	TCAAAAGCTC
	AAGTAGGATG	AAATGGTTAT	ACAAACGGTG	GACCAGAGGA	AGTTTCGAG
4001	GACTTCACCT	CCAAGGGAGG	AGTAATGCCT	GGAGACCTCA	GGTGAATAAT
	CTGAAGTCCA	GGTTCCCTCC	TCATTACGGA	CCTCTGGAGT	CCACTTATTA
4051	CCAAAAGAGT	GGCTGCAAGT	GGACTTCCAG	AAGACAATGA	AAGTCACAGG
	GGTTTCTCA	CCGACGTTCA	CCTGAAGGTC	TTCTGTTACT	TTCAGTGTCC
4101	AGTAACTACT	CAGGGAGTAA	AATCTCTGCT	TACCAGCATG	TATGTGAAGG
	TCATTGATGA	GTCCCTCATT	TTAGAGACGA	ATGGTCGTAC	ATACACTTCC
4151	AGTTCCCTCAT	CTCCAGCAGT	CAAGATGGCC	ATCAGTGGAC	TCTCTTTTT
	TCAAGGAGTA	GAGGTCGTCA	GTTCTACCGG	TAGTCACCTG	AGAGAAAAAA
4201	CAGAATGGCA	AAGTAAAGGT	TTTCAGGGA	AATCAAGACT	CCTTCACACC
	GTCTTACCGT	TTCATTCCA	AAAACTCCCT	TTAGTTCTGA	GGAAGTGTGG
4251	TGTGGTGAAC	TCTCTAGACC	CACCGTTACT	GAATCGCTAC	CTTCGAATTC
	ACACCACTTG	AGAGATCTGG	GTGGCAATGA	CTGAGCGATG	GAAGCTTAAG
4301	ACCCCCAGAG	TTGGGTGCAC	CAGATTGCC	TGAGGATGGA	GGTTCTGGGC
	TGGGGGTCTC	AACCCACGTG	GTCTAACGGG	ACTCCTACCT	CCAAGACCCG
4351	TGCGAGGCAC	AGGACCTTA	C		
	ACGCTCCGTG	TCCTGGAGAT	G		

1-57 SIGNAL PEPTIDE
58-1173 A1 DOMAIN
1174-2277 A2 DOMAIN
2278-2319 SQ LINKER
2320-3432 ap-A3 DOMAINS
3433-3891 C1 DOMAIN
3892-4371 C2 DOMAIN

FIG. 7D

AMINO ACID SEQUENCE OF HP47/OL

1	MQLELSTCVF	LCLLPLGFSA	IRRYYLGAVE	LSWDYRQSEL	LRELHVDTRF
51	PATAPGALPL	GPSVLYKKTV	FVEFTDQLFS	VARPRPPWMG	LLGPTIQAEV
101	YDTVVVTLKN	MASHPVSLHA	VGVSFWKSSE	GAEYEDHTSQ	REKEDDKVLP
151	GKSQTYVWQV	LKENGPTASD	PPCLTYSYLS	HVDLVKDLNS	GLIGALLVCR
201	EGSLTRERTQ	NLHEFVLLFA	VFDEGKSWHS	ARNDSWTRAM	DPAPARAQPA
251	MHTVNGYVNR	SLPGLIGCHK	KSVYWHVIGM	GTSPEVHSIF	LEGHTFLVRH
301	HRQASLEISP	LTFLTAQTFL	MDLGQFLFC	HISSHGGM	EAHVRVESCA
351	EEPQLRRKAD	EEEDYDDNLY	DSDMDVVRLD	GDDVSPFIQI	RSVAKHPKT
401	WVHYIAAEEE	DWDYAPLVLA	PDDRSYKSQY	LNNGPQRIGR	KYKKVRFMAY
451	TDETFKTREA	IQHESGILGP	LLYGEVGDTL	LIIFKNQASR	PYNIYPHGIT
501	DVRPLYSRRL	PKGVKHLKDF	PILPGEIFKY	KWTVTVEDGP	TKSDPRCLTR
551	YYSSFVNMER	DLASGLIGPL	LICYKESVDQ	RGNQIMSDKR	NVILFSVFDE
601	NRSWYLTENI	QRFLPNPAGV	QLEDPEFQAS	NIMHSINGYV	FDSLQLSVCL
651	HEVAYWYIILS	IGAQTDLSV	FFSGYTFKHK	MVYEDTLTLF	PFSGETVFMS
701	MENPGLWILG	CHNSDFRNRG	MTALLKVSSC	DKNTGDYYED	SYEDISAYLL
751	SKNNAIEPRS	FAQNSRPPSA	SAPKPPVLRR	HQRDISLPTF	QPEEDKMDYD
801	DIFSTETKGE	DFDIYGEDEN	QDPRSFQKRT	RHYFIAAVEQ	LWDYGMSESP
851	RALRNRAQNG	EVPRFKVVF	REFADGFTQ	PSYRGELNKH	LGLLGPYIRA
901	EVEDNIMVTF	KNQASRPYSF	YSSLISYPDD	QEQQAEPRHN	FVQPNETRTY
951	FWKVQHHMAP	TEDEFDCKAW	AYFSDVDLEK	DVHSGLIGPL	LICRANTLNA
1001	AHGRQVTVQE	FALFFTIFDE	TKSWYFTENV	ERNCRAPCHL	QMEDPTLKEN
1051	YRFHAINGYV	MDTLPGLVMA	QNQRIRWYLL	SMGSNENIHS	IHFSGHVFST
1101	RKKEEYKMAV	YNLYPGVFET	VEMLPSKVG	WRIECLIGEH	LQAGMSTTFL
1151	VYSKKCQTPL	GMASGHIRDF	QITASGQYQ	WAPKLARLHY	SGSINAWSTK
1201	EPFSWIKVDL	LAPMIIHGIK	TQGARQKFSS	LYISQFIIMY	SLDGKKWQTY
1251	RGNSTGTLMV	FFGNVDSSGI	KHNIFNPPII	ARYIRLHPTH	YSIRSTLRME
1301	LMGCDLNSCS	MPLGMESKAI	SDAQITASSY	FTNMFATWSP	SKARLHLOGR
1351	SNAWRPQVNN	PKEWLQVDFQ	KTMKVTGVTT	QGVKSLLTSM	YVKEFLISSS
1401	QDGHQWTLFF	QNGKVKVFQG	NQDSFTPVNN	SLDPPLLTRY	LRIHPQSWVH
1451	QIALRMEVLG	CEAQDLY*			

1-19 SIGNAL PEPTIDE
 20-391 A1 DOMAIN
 392-759 A2 DOMAIN
 760-783 OL LINKER
 784-1154 ap-A3
 1155-1307 C1 DOMAIN
 1308-1467 C2 DOMAIN

FIG. 8

HP47/OL NUCLEOTIDE SEQUENCE

1 ATGCAGCTAG AGCTCTCCAC CTGTGTCTTT CTGTGTCTCT TGCCACTCGG
51 TACGTCGATC TCGAGAGGTG GACACAGAAA GACACAGAGA ACGGTGAGCC
CTTAGTGCC ATCAGGAGAT ACTACCTGGG CGCAGTGGAA CTGTCCTGGG
GAAATCACGG TAGTCCTCTA TGATGGACCC GCGTCACCTT GACAGGACCC
101 ACTACCAGCA AAGTGAACTC CTCCGTGAGC TGCACGTGGA CACCAAGATT
TGATGGCCGT TTCACTTGAG GAGGCACTCG ACGTGCACCT GTGGTCTAAA
151 CCTGCTACAG CGCCAGGAGC TCTTCCGTT GGCCCCTGAG TCCTGTACAA
GGACGATGTC GCGGTCTCG AGAAGGCAAC CGGGCAGTC AGGACATGTT
201 AAAGACTGTG TTCGTAGAGT TCACGGATCA ACTTTCAAGC GTTGCCAGGC
TTTCTGACAC AAGCATCTCA AGTGCCTAGT TGAAAAGTCG CAACGGTCCG
251 CCAGGCCACC ATGGATGGGT CTGCTGGGT CTACCATCCA GGCTGAGGTT
GGTCCGGTGG TACCTACCCA GACGACCCAG GATGGTAGGT CCGACTCAA
301 TAGGACACGG TGGTCGTTAC CCTGAAGAAC ATGGCTTCTC ATCCCGTTAG
ATGCTGTGCC ACCAGCAATG GGACTTCTTG TACCGAAGAG TAGGGCAATC
351 TCTTCACGCT GTCGGCGTCT CCTTCTGGAA ATCTTCGAA GGCCTGAAT
AGAAGTGCAG CAGCCGCAGA GGAAGACCTT TAGAAGGCTT CCGCAGCTTA
401 ATGAGGATCA CACCAGCCAA AGGGAGAAGG AAGACGATAA AGTCCTTCCC
TACTCCTAGT GTGGTCGGTT TCCCTCTTCC TTCTGCTATT TCAGGAAGGG
451 GGTAAAAGCC AACCTACGT CTGGCAGGTC CTGAAAGAAA ATGGTCCAAC
CCATTTTCCG TTTGGATGCA GACCGTCCAG GACTTCTTT TACCAAGGTTG
501 AGCCTCTGAC CCACCATGTC TTACCTACTC ATACCTGTCT CACGTGGACC
TCGGAGACTG GGTGGTACAG AATGGATGAG TATGGACAGA GTGCACCTGG
551 TGGTGAAGA CCTGAATTG GGCCTCATTG GAGCCCTGCT GTTTGTAGA
ACCACTTTCT GGACTTAAGC CCGGAGTAAC CTCGGGACGA CCAAACATCT
601 GAAGGGAGTC TGACCCAGAGA AAGGACCCAG AACCTGCACG AATTGTACT
CTTCCCTCAG ACTGGTCTCT TTCTGGGTG TTGGACGTGC TTAAACATGA
651 ACTTTTGCT GTCTTGATG AAGGGAAAAG TTGGCACTCA GCAAGAAATG
TGAAAAACGA CAGAAACTAC TTCCCTTTTC AACCCTGAGT CGTTCTTAC
701 ACTCCTGGAC ACGGGCCATG GATCCCGCAC CTGCCAGGGC CCAGCCTGCA
TGAGGACCTG TGCCCCGGTAC CTAGGGCGTG GACGGTCCCC GTGCGGACGT
751 ATGCACACAG TCAATGGCTA TGTCAACAGG TCTCTGCCAG GTCTGATCGG
TACGTGTGTC AGTTACCGAT ACAGTTGTC ACAGACGGTC CAGACTAGCC
801 ATGTCTATAAG AAATCAGTCT ACTGGCACGT GATTGGAATG GGCACCAAGCC
TACAGTATTG TTTAGTCAGA TGACCGTGCA CTAACCTTAC CCGTGGTCGG
851 CGGAAGTGCA CTCCATTG CTTGAAGGCC ACACGTTCT CGTGAGGCAC
GCCCTCACGT GAGGTAAAAA GAACTCCGG TGTGCAAAGA GCACTCCGTG
901 CATGCCAGG CTTCCCTGGA GATCTGCCA CTAACCTTCC TCACTGCTCA
GTAGCGGTCC GAAGGAACCT CTAGAGCGGT GATTGAAAGG AGTACGAGT
951 GACATTCTG ATGGACCTG GCCAGTTCCT ACTGTTTG CATATCTCTT
CTGTAAGGAC TACCTGGAAC CGGTCAAGGA TGACAAAACA GTATAGAGAA
1001 CCCACCACCA TGGTGGCATG GAGGCTCACG TCAGAGTAGA AAGCTGCGCC
GGGTGGTGGT ACCACCGTAC CTCCGAGTGC AGTCTCATCT TTGACGCGG
1051 GAGGAGCCCC AGCTGCGGAG GAAAGCTGAT GAAGAGGAAG ATTATGATGA
CTCCTCGGGG TCGACGCCCT CTTTCGACTA CTTCTCCTTC TAATACTACT
1101 CAATTGTCAC GACTCGGACA TGGACGTGGT CCGGCTCGAT GGTGACGACG
GTTAAACATG CTGAGCCTGT ACCTGCACCA GGCCGAGCTA CCACTGCTGC
1151 TGTCTCCCTT TATCCAAATC CGCTCGGTTG CCAAGAAGCA TCCTAAAAC
ACAGAGGGAA ATAGGTTAG GCGAGCCAAC GGTTCTTCGT AGGATTTGA
1201 TGGGTACATT ACATTGCTGC TGAAGAGGAG GACTGGGACT ATGCTCCCTT

FIG. 9A

21/33

1251	ACCCATGTAA	TGTAACGACG	ACTTCTCCTC	CTGACCCTGA	TACGAGGGAA
	AGTCCTCGCC	CCCGATGACA	GAAGTTATAA	AAGTCAATAT	TTGAACAATG
1301	TCAGGAGCGG	GGGCTACTGT	CTTCAATATT	TTCAGTTATA	AACTTGTAC
	GCCCTCAGCG	GATTGGTAGG	AAGTACAAAA	AAGTCGATT	TATGGCATAAC
1351	CGGGAGTCGC	CTAACCATCC	TTCATGTTT	TTCAGGCTAA	ATACCCTATG
	ACAGATGAAA	CCTTAAGAC	GCGTGAAGCT	ATTCAAGCATG	AATCAGGAAT
1401	TGTCTACTTT	GGAAATTCTG	CGCACTTCGA	TAAGTCGTAC	TTAGTCCTTA
	CTTGGGACCT	TTACTTTATG	GGGAAGTTGG	AGACACACTG	TTGATTATAT
1451	GAACCTGGA	AATGAAATAC	CCCTCAACC	TCTACCCCTCA	CGGAATCACT
	TTAAGAATCA	AGCAAGCAGA	CCATATAACA	GGTTTTCCAC	ATTTGTAAA
1501	AATTCTTAGT	TCGTTCGTCT	GGTATATTGT	AGATGGGAGT	GCCTTAGTGA
	GATGTCCGTC	CTTGTATTTC	AAGGAGATTA	CCAAAAGGTG	TAAAACATT
1551	CTACAGGCAG	GAAACATAAG	TTCCCTCTAA	GGTTTTCCAC	ATTTGTAAA
	GAAGGATTTT	CCAATTCTGC	CAGGAGAAAT	ATTCAAATAT	AAATGGACAG
1601	CTTCCTAAAA	GGTTAAGACG	GTCCTCTTAA	TAAGTTTATA	TTTACCTGTC
	TGACTGTAGA	AGATGGGCCA	ACTAAATCAG	ATCCCGGGTG	CCTGACCCGC
1651	ACTGACATCT	TCTACCCGGT	TGATTTAGTC	TAGGCGCCAC	GGACTGGCG
	TATTACTCTA	GTTTCGTTAA	TATGGAGAGA	GATCTAGCTT	CAGGACTCAT
1701	ATAATGAGAT	CAAAGCAATT	ATACCTCTCT	CTAGATCGAA	GTCTGAGTA
	TGGCCCTCTC	CTCATCTGCT	ACAAAGAAC	TGTAGATCAA	AGAGGAAACC
1751	ACCGGGAGAG	GAGTAGACGA	TGTTTCTTAG	ACATCTAGTT	TCTCCTTTGG
	AGATAATGTC	AGACAAGAGG	AATGTCATCC	TGTTTTCTGT	ATTGATGAG
1801	TCTATTACAG	TCTGTTCTCC	TTACAGTAGG	ACAAAAGACA	TAAACTACTC
	AACCAGAGCT	GGTACCTCAC	AGAGAATATA	CAACGCTTTC	TCCCCAATCC
1851	TTGGCTTCGA	CCATGGAGTG	TCTCTTATAT	GTTGCAGAAAG	AGGGGTTAGG
	AGCTGGAGTG	CAGCTTGAGG	ATCCAGAGTT	CCAAGCCTCC	AACATCATGC
1901	TCGACCTCAC	GTCGAACCTC	TAGGTCTCAA	GGTTCGGAGG	TTGTAGTACG
	ACAGCATCAA	TGGCTATGTT	TTTGATAGTT	TGCAGTTGTC	AGTTTGTGG
1951	TGTCGTAGTT	ACCGATACAA	AAACTATCAA	ACGTCAACAG	TCAAACAAAC
	CATGAGGTGG	CATACTGGTA	CATTCTAAC	ATTGGAGCAC	AGACTGACTT
2001	GTACTCCACC	GTATGACCAT	GTAAGATTG	TAACCTCGTG	TCTGACTGAA
	CCTTTCTGTC	TTCTTCTCTG	GATATACCTT	CAAACACAAA	ATGGTCTATG
2051	GGAAAGACAG	AAGAAGAGAC	CTATATGGAA	GTTGTGTTT	TACCAAGATA
	AAGACACACT	CACCCATTTC	CCATTCTCAG	GAGAAACTGT	CTTCATGTCG
2101	TTCTGTGTGA	GTGGGATAAG	GGTAAGAGTC	CTCTTGACA	GAAGTACAGC
	ATGGAAAACC	CAGGTCTATG	GATTCTGGGG	TGCCACAAC	CAGACTTCG
2151	TACCTTTGG	GTCCAGATAC	CTAACGACCC	ACGGTGTGA	GTCTGAAAGC
	GAACAGAGGC	ATGACCGCCT	TACTGAAGGT	TTCTAGTTGT	GACAAGAAC
2201	CTTGTCTCCG	TACTGGCGGA	ATGACTTCCA	AAGATCAACA	CTGTTCTGT
	CTGGTGATTA	TTACGAGGAC	AGTTATGAAG	ATATTCAGC	ATACTGCTG
2251	GACCACTAAT	TTACGAGGAC	TCAATACCTC	TATAAAGTCG	TATGAACGAC
	AGTAAAACA	ATGCCATTGA	ACCTAGGAGC	TTTGCCCAGA	ATTCAAGACC
2301	TCATTTTGT	TACGGTAACT	TGGATCCTCG	AAACGGGTCT	TAAGTCTGG
	CCCTAGTGGC	AGCGCTCCAA	AGCCTCCGGT	CCTGCGACGG	CATCAGAGGG
2351	GGGATCACCG	TCGCGAGGTT	TCGGAGGCCA	GGACGCTGCC	GTAGTCTCCC
	ACATAAGCCT	TCCTACTTT	CAGCCGGAGG	AAGACAAAAT	GGACTATGAT
2401	TGTATTCGGA	AGGATGAAAA	GTCGGCCTCC	TTCTGTTTA	CCTGATACTA
	GATATCTTCT	CAACTGAAAC	GAAGGGAGAA	GATTTGACA	TTTACGGTGA
2451	CTATAGAAGA	GTTGACTTTG	CTTCCCTCTT	CTAAAACGT	AAATGCCACT
	GGATGAAAAT	CAGGACCCCTC	GCAGCTTTCA	GAAGAGAAC	CGACACTATT
2501	CCTACTTTA	GTCCCTGGGAG	CGTCGAAAGT	CTTCTCTTGG	GCTGTGATAAA
	TCATTGCTGC	GGTGGAGCAG	CTCTGGGATT	ACGGGATGAG	CGAATCCCCC

FIG. 9B

2551	AGTAACGACG	CCACCTCGTC	GAGACCCTAA	TGCCCTACTC	GCTTAGGGGG
	CGGGCGCTAA	GAAACAGGGC	TCAGAACGGA	GAGGTGCCTC	GGTTCAAGAA
2601	GCCCACGATT	CTTGTCCTCG	AGTCTTGCT	CTCCACGGAG	CCAAGTTCTT
	GGTGGTCTTC	CGGGAATTG	CTGACGGCTC	CTTCACGAG	CCGTCGTACC
2651	CCACCCAGAAG	GCCCTTAAAC	GACTGCCGAG	GAAGTGCCTC	GGCAGCATGG
	GCGGGGAACT	CAACAAACAC	TTGGGGCTCT	TGGGACCCCTA	CATCAGAGCG
2701	CGCCCCTTGA	GTTGTTTGTG	AACCCCGAGA	ACCCTGGGAT	GTAGTCTCGC
	GAAGTTGAAG	ACAACATCAT	GGTAACCTTC	AAAAACCAGG	CGTCTCGTCC
2751	CTTCAACTTC	TGTTGTAGTA	CCATTGAAAG	TTTTGGTCC	GCAGAGCAGG
	CTATTCTTC	TACTCGAGCC	TTATTTCTTA	TCCGGATGAT	CAGGAGCAAG
2801	GATAAGGAAG	ATGAGCTCGG	AATAAAGAAT	AGGCCTACTA	GTCCCTCGTTC
	GGGCAGAACCC	TCGACACAAAC	TTCGTCCAGC	CAAATGAAAC	CAGAACTTAC
2851	CCCCTCTTGG	AGCTGTGTTG	AAGCAGGTCTG	GTAACTTTG	GTCTTGAATG
	TTTTGGAAAG	TGCAGCATCA	CATGGCACCC	ACAGAAGACG	AGTTTGACTG
2901	AAAACCTTTC	ACGTCGTAGT	GTACCGTGGG	TGTCTTCTGC	TCAAACGTGAC
	CAAAGCCTGG	GCCTACTTTT	CTGATGTTGA	CCTGGAAAAAA	GATGTGCACT
2951	GTTCGGGACC	CGGATGAAAAA	GACTACAAC	GGACCTTTT	CTACACGTGA
	CAGGCTTGTAT	CGGCCCCCTT	CTGATCTGCC	GCGCCAACAC	CCTGAACGCT
3001	GTCCGAACTA	GCCGGGGGAA	GAATAGACGG	CGCGGTTGTG	GGACTTGCAG
	GCTCACGGTA	GACAAGTGAC	CGTGCAAGAA	TTTGCTCTGT	TTTCACTAT
3051	CGAGTGCAT	CTGTTCACTG	GCACGTTCTT	AAACGAGACA	AAAAGTATA
	TTTGATGAG	ACAAAGAGCT	GGTACTTCAC	TGAAAATGTG	GAAAGGAACT
3101	AAAACACTTC	TGTTTCTCGA	CCATGAAGTG	ACTTTACAC	CTTCCCTTGA
	GCCGGGCCCC	CTGCCATCTG	CAGATGGAGG	ACCCCACCT	GAAAGAAAAC
3151	CGGCCCGGGG	GACGGTAGAC	GTCTACCTCC	TGGGGTGAGA	CTTCTTTTG
	TATCGCTTCC	ATGCAATCAA	TGGCTATGTG	ATGGATACAC	TCCCTGGCTT
3201	ATAGCGAAGG	TACGTTAGTT	ACCGATACAC	TACCTATGTG	AGGGACCGAA
	AGTAATGGCT	CAGAATCAA	GGATCCGATG	GTATCTGCTC	AGCATGGGCA
3251	TCATTACCGA	GTCTTAGTTT	CCTAGGCTAC	CATAGACGAG	TCGTACCCGT
	GCAATGAAAAA	TATCCATTCTG	ATTCAATTAA	GCGGACACGT	GTTCAGTGT
3301	CGTTACTTTT	ATAGGTAAGC	TAAGTAAAT	CGCCTGTGCA	CAAGTCACAT
	CGGAAAAAAGG	AGGAGTATAA	AATGGCCGTG	TACAATCTCT	ATCCGGGTGT
3351	GCCTTTTCC	TCCTCATATT	TTACCGGCAC	ATGTTAGAGA	TAGGCCACAA
	CTTGAGACA	GTGGAAATGC	TACCGTCAA	AGTTGGAATT	TGGCGAATAG
3401	GAAACTCTGT	CACCTTACG	ATGGCAGGTT	TCAACCTTAA	ACCGCTTATC
	AATGCCTGAT	TGGCGAGCAC	CTGCAAGCTG	GGATGAGCAC	GACTTTCTG
3451	TTACGGACTA	ACCGCTCGTG	GACGTTCGAC	CCTACTCGTG	CTGAAAGGAC
	GTGTACAGCA	AGAAGTGTCA	GACTCCCCTG	GGAATGGCTT	CTGGACACAT
3501	CACATGTCGT	TCTTCACAGT	CTGAGGGGAC	CCTTACCGAA	GACCTGTGTA
	TAGAGATTAA	CAGATTACAG	CTTCAGGACA	ATATGGACAG	TGGGCCACAA
3551	ATCTCTAAAA	GTCTAATGTC	GAAGTCTGT	TATACCTGTC	ACCCGGGGTT
	AGCTGGCCAG	ACTTCATTAT	TCCGGATCAA	TCAATGCCTG	GAGCACCAAG
3601	TCGACCGGTC	TGAAGTAATA	AGGCCTAGTT	AGTTACGGAC	CTCGTGGTCA
	GAGCCCTTTT	CTTGGATCAA	GGTGGATCTG	TTGGCACCAA	TGATTATTCA
3651	CTCGGGAAAAA	GAACCTAGTT	CCACCTAGAC	AACCCTGGTT	ACTAATAAGT
	CGGCATCAAG	ACCCAGGGTG	CCCGTCAGAA	GTTCTCCAGC	CTCTACATCT
3701	GCCGTAGTTC	TGGGTCCCAC	GGGCAGTCTT	CAAGAGGTCTG	GAGATGTAGA
	CTCAGTTTAT	CATCATGTAT	AGTCTTGATG	GGAAGAAGTG	GCAGACTTAT
3751	GAGTCAAATA	GTAGTACATA	TCAGAACTAC	CCTTCTTCAC	CGTCTGAATA
	CGAGGAAATT	CCACTGGAAC	CTTAATGGTC	TTCTTTGGCA	ATGTGGATTC
3801	GCTCCTTAA	GGTGACCTTG	GAATTACCAAG	AAGAAACCGT	TACACCTAAG
	ATCTGGGATA	AAACACAATA	TTTTTAACCC	TCCAATTATT	GCTCGATACA

23/33

3851	TAGACCCTAT	TTTGTGTTAT	AAAAATTGGG	AGGTTAATAA	CGAGCTATGT
	TCCGTTTGCA	CCCAACTCAT	TATAGCATTC	GCAGCACTCT	TCGCATGGAG
3901	AGGCAAACGT	GGGTTGAGTA	ATATCGTAAG	CGTCGTGAGA	AGCGTACCTC
	TTGATGGGCT	GTGATTAA	TAGTTGCAGC	ATGCCATTGG	GAATGGAGAG
3951	AACTACCCGA	CACTAAATT	ATCAACGTCG	TACGGTAACC	CTTACCTCTC
	TAAAGCAATA	TCAGATGCAC	AGATTACTGC	TTCATCCTAC	TTTACCAATA
	ATTTCGTTAT	AGTCTACGTG	TCTAATGACG	AAGTAGGATG	AAATGGTTAT
4001	TGTTTGCAC	CTGGTCTCCT	TCAAAAGCTC	GACTTCACCT	CCAAGGGAGG
	ACAAACGGTG	GACCAGAGGA	AGTTTCGAG	CTGAAGTGGA	GGTCCCTCC
4051	AGTAATGCCT	GGAGACCTCA	GGTGAATAAT	CCAAAAGAGT	GGCTGCAAGT
	TCATTACGGA	CCTCTGGAGT	CCACTTATTA	GGTTTCTCA	CCGACGTTCA
4101	GGACTTCCAG	AAGACAAATGA	AAGTCACAGG	AGTAACTACT	CAGGGAGTAA
	CCTGAAGGTC	TTCTGTTACT	TTCAGTGTCC	TCATTGATGA	GTCCCTCATT
4151	AATCTCTGCT	TACCAGCATG	TATGTGAAGG	AGTTCCCTCAT	CTCCAGCAGT
	TTAGAGACGA	ATGGTCGTAC	ATACACTTCC	TCAAGGAGTA	GAGGTCGTCA
4201	CAAGATGGCC	ATCAGTGGAC	TCTCTTTTT	CAGAATGGCA	AAGTAAAGGT
	GTTCTACCGG	TAGTCACCTG	AGAGAAAAAA	GTCTTACCGT	TTCATTTCCA
4251	TTTCAGGGA	AATCAAGACT	CCTTCACACC	TGTGGTGAAC	TCTCTAGACC
	AAAAGTCCCT	TTAGTTCTGA	GGAAGTGTGG	ACACCACTG	AGAGATCTGG
4301	CACCGTTACT	GACTCGCTAC	CTTCGAATTC	ACCCCCAGAG	TTGGGTGCAC
	GTGGCAATGA	CTGAGCGATG	GAAGCTTAAG	TGGGGGTCTC	AACCCACGTG
4351	CAGATTGCC	TGAGGATGGA	GGTTCTGGC	TGCGAGGCAC	AGGACCTCTA
	GTCTAACGGG	ACTCCTACCT	CCAAGACCCG	ACGCTCCGTG	TCCTGGAGAT
4401	C				
	G				

FIG. 9D

AMINO ACID SEQUENCE OF HUMAN B DOMAIN-DELETED FACTOR VIII (HSQ)

Met Gln Ile Glu Leu Ser Thr Cys Phe Phe Leu Cys Leu Leu Arg Phe
1 5 10 15
Cys Phe Ser Ala Thr Arg Arg Tyr Tyr Leu Gly Ala Val Glu Leu Ser
20 25 30
Trp Asp Tyr Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg
35 40 45
Phe Pro Pro Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val
50 55 60
Tyr Lys Lys Thr Leu Phe Val Glu Phe Thr Val His Leu Phe Asn Ile
65 70 75 80
Ala Lys Pro Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile Gln
85 90 95
Ala Glu Val Tyr Asp Tyr Val Val Ile Thr Leu Lys Asn Met Ala Ser
100 105 110
His Pro Val Ser Leu His Ala Val Gly Val Ser Tyr Trp Lys Ala Ser
115 120 125
Glu Gly Ala Glu Thr Asp Asp Gln Thr Ser Gln Arg Glu Lys Glu Asp
130 135 140
Asp Lys Val Phe Pro Gly Gly Ser His Thr Tyr Val Trp Gln Val Leu
145 150 155 160
Lys Glu Asn Gly Pro Met Ala Ser Asp Pro Leu Cys Leu Thr Tyr Ser
165 170 175
Tyr Leu Ser His Val Asp Leu Val Lys Asp Leu Asn Ser Gly Leu Ile
180 185 190
Gly Ala Leu Leu Val Cys Arg Glu Gly Ser Leu Ala Lys Glu Lys Thr
195 200 205
Gln Thr Leu His Lys Phe Ile Leu Leu Phe Ala Val Phe Asp Glu Gly
210 215 220
Lys Ser Trp His Ser Glu Thr Lys Asn Ser Leu Met Gln Asp Arg Asp
225 230 235 240
Ala Ala Ser Ala Arg Ala Trp Pro Lys Met His Thr Val Asn Gly Tyr
245 250 255
Val Asn Arg Ser Leu Pro Gly Leu Ile Gly Cys His Arg Lys Ser Val
260 265 270
Tyr Trp His Val Ile Gly Met Gly Thr Thr Pro Glu Val His Ser Ile
275 280 285
Phe Leu Glu Gly His Thr Phe Leu Val Arg Asn His Arg Gln Ala Ser
290 295 300
Leu Glu Ile Ser Pro Ile Thr Phe Leu Thr Ala Gln Thr Leu Leu Met
305 310 315 320
Asp Leu Gly Gln Phe Leu Leu Phe Cys His Ile Ser Ser His Gln His
325 330 335
Asp Gly Met Glu Ala Tyr Val Lys Val Asp Ser Cys Pro Glu Glu Pro
340 345 350
Gln Leu Arg Met Lys Asn Asn Glu Glu Ala Glu Asp Tyr Asp Asp Asp
355 360 365
Leu Thr Asp Ser Glu Met Asp Val Val Arg Phe Asp Asp Asp Asn Ser
370 375 380
Pro Ser Phe Ile Gln Ile Arg Ser Val Ala Lys Lys His Pro Lys Thr
385 390 395 400
Trp Val His Tyr Ile Ala Ala Glu Glu Asp Trp Asp Tyr Ala Pro

25/33

	405	410	415													
Leu	Val	Leu	Ala	Pro	Asp	Asp	Arg	Ser	Tyr	Lys	Ser	Gln	Tyr	Leu	Asn	
				420				425					430			
Asn	Gly	Pro	Gln	Arg	Ile	Gly	Arg	Lys	Tyr	Lys	Lys	Val	Arg	Phe	Met	
				435				440				445				
Ala	Tyr	Thr	Asp	Glu	Thr	Phe	Lys	Thr	Arg	Glu	Ala	Ile	Gln	His	Glu	
				450				455			460					
Ser	Gly	Ile	Leu	Gly	Pro	Leu	Leu	Tyr	Gly	Glu	Val	Gly	Asp	Thr	Leu	
				465				470			475			480		
Leu	Ile	Ile	Phe	Lys	Asn	Gln	Ala	Ser	Arg	Pro	Tyr	Asn	Ile	Tyr	Pro	
				485				490			495					
His	Gly	Ile	Thr	Asp	Val	Arg	Pro	Leu	Tyr	Ser	Arg	Arg	Leu	Pro	Lys	
				500				505				510				
Gly	Val	Lys	His	Leu	Lys	Asp	Phe	Pro	Ile	Leu	Pro	Gly	Glu	Ile	Phe	
				515				520				525				
Lys	Tyr	Lys	Trp	Thr	Val	Thr	Val	Glu	Asp	Gly	Pro	Thr	Lys	Ser	Asp	
				530				535				540				
Pro	Arg	Cys	Leu	Thr	Arg	Tyr	Tyr	Ser	Ser	Phe	V ^a l	Asn	Met	Glu	Arg	
				545				550				555			560	
Asp	Leu	Ala	Ser	Gly	Leu	Ile	Gly	Pro	Leu	Leu	Ile	Cys	Tyr	Lys	Glu	
				565				570				575				
Ser	Val	Asp	Gln	Arg	Gly	Asn	Gln	Ile	Met	Ser	Asp	Lys	Arg	Asn	Val	
				580				585				590				
Ile	Leu	Phe	Ser	Val	Phe	Asp	Glu	Asn	Arg	Ser	Trp	Tyr	Leu	Thr	Glu	
				595				600				605				
Asn	Ile	Gln	Arg	Phe	Leu	Pro	Asn	Pro	Ala	Gly	Val	Gln	Leu	Glu	Asp	
				610				615				620				
Pro	Glu	Phe	Gln	Ala	Ser	Asn	Ile	Met	His	Ser	Ile	Asn	Gly	Tyr	Val	
				625				630				635			640	
Phe	Asp	Ser	Leu	Gln	Leu	Ser	Val	Cys	Leu	His	Glu	Val	Ala	Tyr	Trp	
				645				650				655				
Tyr	Ile	Leu	Ser	Ile	Gly	Ala	Gln	Thr	Asp	Phe	Leu	Ser	Val	Phe	Phe	
				660				665				670				
Ser	Gly	Tyr	Thr	Phe	Lys	His	Lys	Met	Val	Tyr	Glu	Asp	Thr	Leu	Thr	
				675				680				685				
Leu	Phe	Pro	Phe	Ser	Gly	Glu	Thr	Val	Phe	Met	Ser	Met	Glu	Asn	Pro	
				690				695				700				
Gly	Leu	Trp	Ile	Leu	Gly	Cys	His	Asn	Ser	Asp	Phe	Arg	Asn	Arg	Gly	
				705				710				715			720	
Met	Thr	Ala	Leu	Leu	Lys	Val	Ser	Ser	Cys	Asp	Lys	Asn	Thr	Gly	Asp	
				725				730				735				
Tyr	Tyr	Glu	Asp	Ser	Tyr	Glu	Asp	Ile	Ser	Ala	Tyr	Leu	Leu	Ser	Lys	
				740				745				750				
Asn	Asn	Ala	Ile	Glu	Pro	Arg	Ser	Phe	Ser	Gln	Asn	Pro	Pro	Val	Leu	
				755				760				765				
Lys	Arg	His	Gln	Arg	Glu	Ile	Thr	Arg	Thr	Thr	Leu	Gln	Ser	Asp	Gln	
				770				775				780				
Glu	Glu	Ile	Asp	Tyr	Asp	Asp	Thr	Ile	Ser	Val	Glu	Met	Lys	Lys	Glu	
				785				790				795			800	
Asp	Phe	Asp	Ile	Tyr	Asp	Glu	Asp	Glu	Asn	Gln	Ser	Pro	Arg	Ser	Phe	
				805				810				815				
Gln	Lys	Lys	Thr	Arg	His	Hyd	Phe	Ile	Ala	Ala	Val	Glu	Arg	Leu	Trp	
				820				825				830				
Asp	Tyr	Gly	Met	Ser	Ser	Ser	Pro	His	Val	Leu	Arg	Asn	Atg	Ala	Gln	
				835				840				845				
Ser	Gly	Ser	Val	Pro	Gln	Phe	Lys	Lys	Val	Val	Phe	Gln	Glu	Phe	Thr	
				850				855				860				

FIG. 10B

26/33

Asp Gly Ser Phe Thr Gln Pro Leu Tyr Arg Gly Glu Leu Asn Glu His
865 870 875 880
Leu Gly Leu Leu Gly Pro Tyr Ile Arg Ala Glu Val Glu Asp Asn Ile
885 890 895
Met Val Thr Phe Arg Asn Gln Ala Ser Arg Pro Tyr Ser Phe Tyr Ser
900 905 910
Ser Leu Ile Ser Tyr Glu Glu Asp Gln Arg Gln Gly Ala Glu Pro Arg
915 920 925
Lys Asn Phe Val Lys Pro Asn Glu Thr Lys Thr Tyr Phe Trp Lys Val
930 935 940
Gln His His Met Ala Pro Thr Lys Asp Glu Phe Asp Cys Lys Ala Trp
945 950 955 960
Ala Tyr Phe Ser Asp Val Asp Leu Glu Lys Asp Val His Ser Gly Leu
965 970 975
Ile Gly Pro Leu Leu Val Cys His Thr Asn Thr Leu Asn Pro Ala His
980 985 990
Gly Arg Gln Val Thr Val Gln Glu Phe Ala Leu Phe Phe Thr Ile Phe
995 1000 1005
Asp Glu Thr Lys Ser Trp Tyr Phe Thr Glu Asn Met Glu Arg Asn Cys
1010 1015 1020
Arg Ala Pro Cys Asn Ile Gln Met Glu Asp Pro Thr Phe Lys Glu Asn
1025 1030 1035 1040
Tyr Arg Phe His Ala Ile Asn Gly Tyr Ile Met Asp Thr Leu Pro Gly
1045 1050 1055
Leu Val Met Ala Gln Asp Gln Arg Ile Arg Trp Tyr Leu Leu Ser Met
1060 1065 1070
Gly Ser Asn Glu Asn Ile His Ser Ile His Phe Ser Gly His Val Phe
1075 1080 1085
Thr Val Arg Lys Lys Glu Glu Tyr Lys Met Ala Leu Tyr Asn Leu Tyr
1090 1095 1100
Pro Gly Val Phe Glu Thr Val Glu Met Leu Pro Ser Lys Ala Gly Ile
1105 1110 1115 1120
Trp Arg Val Glu Cys Leu Ile Gly Glu His Leu His Ala Gly Met Ser
1125 1130 1135
Thr Leu Phe Leu Val Tyr Ser Asn Lys Cys Gln Thr Pro Leu Gly Met
1140 1145 1150
Ala Ser Gly His Ile Arg Asp Phe Gln Ile Thr Ala Ser Gly Gln Tyr
1155 1160 1165
Gly Gln Trp Ala Pro Lys Leu Ala Arg Leu His Tyr Ser Gly Ser Ile
1170 1175 1180
Asn Ala Trp Ser Thr Lys Glu Pro Phe Ser Trp Ile Lys Val Asp Leu
1185 1190 1195 1200
Leu Ala Pro Met Ile Ile His Gly Ile Lys Thr Gln Gly Ala Arg Gln
1205 1210 1215
Lys Phe Ser Ser Leu Tyr Ile Ser Gln Phe Ile Ile Met Tyr Ser Leu
1220 1225 1230
Asp Gly Lys Lys Trp Gln Thr Tyr Arg Gly Asn Ser Thr Gly Thr Leu
1235 1240 1245
Met Val Phe Phe Gly Asn Val Asp Ser Ser Gly Ile Lys His Asn Ile
1250 1255 1260
Phe Asn Pro Pro Ile Ile Ala Arg Tyr Ile Arg Leu His Pro Thr His
1265 1270 1275 1280
Tyr Ser Ile Arg Ser Thr Leu Arg Met Glu Leu Met Gly Cys Asp Leu
1285 1290 1295
Asn Ser Cys Ser Met Pro Leu Gly Met Glu Ser Lys Ala Ile Ser Asp
1300 1305 1310
Ala Gln Ile Thr Ala Ser Ser Tyr Phe Thr Asn Met Phe Ala Thr Trp

FIG. 10C

27/33

1315 1320 1325
Ser Pro Ser Lys Ala Arg Leu His Leu Gln Gly Arg Ser Asn Ala Trp
1330 1335 1340
Arg Pro Gln Val Asn Asn Pro Lys Glu Trp Leu Gln Val Asp Phe Gln
1345 1350 1355 1360
Lys Thr Met Lys Val Thr Gly Val Thr Thr Gln Gly Val Lys Ser Leu
1365 1370 1375
Leu Thr Ser Met Tyr Val Lys Glu Phe Leu Ile Ser Ser Ser Gln Asp
1380 1385 1390
Gly His Gln Trp Thr Leu Phe Phe Gln Asn Gly Lys Val Lys Val Phe
1395 1400 1405
Gln Gly Asn Gln Asp Ser Phe Thr Pro Val Val Asn Ser Leu Asp Pro
1410 1415 1420
Pro Leu Leu Thr Arg Tyr Leu Arg Ile His Pro Gln Ser Trp Val His
1425 1430 1435 1440
Gln Ile Ala Leu Arg Met Glu Val Leu Gly Cys Glu Ala Gln Asp Leu
1445 1450 1455
Tyr

FIG. 10D

NUCLEOTIDE SEQUENCE OF HUMAN B DOMAIN-DELETED FACTOR VIII (HSQ)

1	ATGCAAATAG	AGCTCTCCAC	CTGCTTCTTT	CTGTGCCTT	TGCGATTCTG
51	CTTTAGTGCC	ACCAGAAGAT	ACTACCTGGG	TGCAGTGGAA	CTGTCATGGG
101	ACTATATGCA	AAGTGATCTC	GGTGAGCTGC	CTGTGGACGC	AAGATTTCT
151	CCTAGAGTGC	CAAAATCTT	TCCATTCAAC	ACCTCAGTCG	TGTACAAAAA
201	GAECTCTGTT	GTAGAATTCA	CGGTTCACCT	TTTCAACATC	GCTAAGCCAA
251	GGCCACCCCTG	GATGGGTCTG	CTAGGTCCTA	CCATCCAGGC	TGAGGTTTAT
301	GATACAGTGG	TCATTACACT	TAAGAACATG	GCTTCCCAC	CTGTCAGTCT
351	TCATGCTGTT	GGTGTATCCT	ACTGGAAAGC	TTCTGAGGGAA	GCTGAATATG
401	ATGATCAGAC	CAGTCAAAGG	GAGAAAGAAG	ATGATAAAAGT	CTTCCCTGGT
451	GGAAAGCCATA	CATATGTCTG	GCAGGGTCTG	AAAGAGAAATG	CTCCAATGGC
501	CTCTGACCCA	CTGTGCCTA	CCTACTCATA	TCTTCTCAT	GTGGACCTGG
551	TAAAAGACTT	GAATTCAAGG	CTCATTGGAG	CCCTACTAGT	ATGTAGAGAA
601	GGGAGTCTGG	CCAAGGAAAA	GACACAGACC	TTGCACAAAT	TTATACTACT
651	TTTGCTGTA	TTTGATGAAG	GGAAAAGTTG	GCACTCAGAA	ACAAAGAACT
701	CCTTGATGCA	GGATAGGGAT	GCTGCATCTG	CTCGGGCCTG	GCCTAAAATG
751	CACACAGTCA	ATGGTTATGT	AAACAGGTCT	CTGCCAGGTC	TGATTGGATG
801	CCACAGGAAA	TCAGTCTATT	GGCATGTGAT	TGGAATGGGC	ACCAACTCCTG
851	AAGTGCACTC	AATATTCCCTC	GAAGGTCACA	CATTCTTGT	GAGGAACCAT
901	CGCCAGGCGT	CCTTGGAAAT	CTCGCCAATA	ACTTTCTTA	CTGCTCAAAC
951	ACTCTTGATG	GACCTTGGAC	AGTTTCTACT	GTTTGTAT	ATCTCTTCCC
1001	ACCAACATGA	TGGCATGGAA	GCTTATGTCA	AAGTAGACAG	CTGTCAGAG
1051	GAACCCCCAAC	TACGAATGAA	AAATAATGAA	GAAGCGGAAG	ACTATGATGA
1101	TGATCTTACT	GATTCTGAAA	TGGATGTGGT	CAGGTTTGAT	GATGACAAC
1151	CTCCTTCCTT	TATCCAAATT	CGCTCAGTTG	CCAAGAAGCA	TCCTAAAAC
1201	TGGGTACATT	ACATTGCTGC	TGAAGAGGAG	GACTGGGACT	ATGCTCCCTT
1251	AGTCCTCGCC	CCCGATGACA	GAAGTTATAA	AAGTCAATAT	TTGAACAAATG
1301	GCCCTCAGCG	GATTGGTAGG	AAGTACAAAAA	AAGTCCGATT	TATGGCATAAC
1351	ACAGATGAAA	CCTTTAAGAC	GCGTGAAGCT	ATTCAAGCATG	AATCAGGAAT
1401	CTTGGGACCT	TTACTTTATG	GGGAAGTTGG	AGACACACTG	TTGATTATAT
1451	TTAAGAATCA	AGCAAGCAGA	CCATATAACA	TCTACCCCTCA	CGGAATCACT
1501	GATGTCCGTC	CTTTGTATTC	AAGGAGATTA	CCAAAAGGTG	TAAAACATTT
1551	GAAGGATTAA	CCAATTCTGC	CAGGAGAAAT	ATTCAAATAT	AAATGGACAG
1601	TGACTGTAGA	AGATGGGCCA	ACTAAATCAG	ATCCCGGGTG	CCTGACCCGC
1651	TATTACTCTA	GTTCGTTAA	TATGGAGAGA	GATCTAGCTT	CAGGACTCAT
1701	TGGCCCTCTC	CTCATCTGCT	ACAAAGAAC	TGTAGATCAA	AGAGGAAACC
1751	AGATAATGTC	AGACAAGAGG	AATGTCATCC	TGTTTCTGT	ATTGATGAG
1801	AACCGAAGCT	GGTACCTCAC	AGAGAATATA	CAACGCTTTC	TCCCCAATCC
1851	AGCTGGAGTG	CAGCTTGAGG	ATCCAGAGTT	CCAAGCCTCC	AACATCATGC
1901	ACAGCATCAA	TGGCTATGTT	TTTGATAGTT	TGCAGTTGTC	AGTTGTTTG
1951	CATGAGGTGG	CATACTGGTA	CATTCTAAC	ATTGGAGCAC	AGACTGACTT
2001	CCTTCTGTC	TTCTTCTCTG	GATATACCTT	CAAACACAAA	ATGGTCTATG
2051	AAGACACACT	CACCCATTTC	CCATTCTCAG	GAGAAACTGT	CTTCATGTCG
2100	ATGGAAAACC	CAGGTCTATG	GATTCTGGGG	TGCCACAAC	CAGACTTCG
2151	GAACAGAGGC	ATGACCGCCT	TACTGAAGGT	TTCTAGTTGT	GACAAGAAC
2201	CTGGTGATTA	TTACCGAGGAC	AGTTATGAAG	ATATTTCAGC	ATACTTGCTG
2251	AGTAAAACAA	ATGCCATTGA	ACCTAGGAGC	TTCTCTCAGA	ATCCACCAAGT
2301	CTTGAACACGC	CATCAACGGG	AAATAACTCG	TACTACTCTT	CAGTCAGATC
2351	AAGAGGAAAT	TGACTATGAT	GATACCATAT	CAGTTGAAAT	GAAGAAGGAA
2401	GATTTGACA	TTTATGATGA	GGATGAAAAT	CAGAGCCCCC	GCAGCTTCA

2451	AAAGAAAACA	CGACACTATT	TTATTGCTGC	AGTGGAGAGG	CTCTGGGATT
2501	ATGGGATGAG	TAGCTCCCCA	CATGTTCTAA	GAAACAGGGC	TCAGAGTGGC
2551	AGTGTCCCTC	AGTTCAAGAA	AGTTGTTTTC	CAGGAATTAA	CTGATGGCTC
2601	CTTACTCAG	CCCTTATACC	GTGGAGAACT	AAATGAACAT	TTGGGACTCC
2651	TGGGCCATA	TATAAGAGCA	GAAGTTGAAG	ATAATATCAT	GGTAACCTTC
2701	AGAAATCAGG	CCTCTCGTCC	CTATTCCCTC	TATTCTAGCC	TTATTCTTA
2751	TGAGGAAGAT	CAGAGGCAAG	GAGCAGAAC	TAGAAAAAAAC	TTTGTCAAGC
2801	CTAATGAAAC	CAAAACTTAC	TTTGGAAAG	TGCAACATCA	TATGGCACCC
2851	ACTAAAGATG	AGTTGACTG	CAAAGCCTGG	GCTTATTCT	CTGATGTTGA
2901	CCTGGAAAAA	GATGTGCACT	CAGGCCTGAT	TGGACCCCTT	CTGGTCTGCC
2951	ACACTAACAC	ACTGAACCCCT	GCTCATGGGA	GACAAGTGC	AGTACAGGAA
3001	TTTGCTCTGT	TTTCACCAT	CTTGATGAG	ACCAAAAGCT	GGTACTTCAC
3051	TGAAAATATG	GAAAGAAACT	GCAGGGCTCC	CTGCAATATC	CAGATGGAAG
3101	ATCCCACCTT	TAAAGAGAAAT	TATCGCTTCC	ATGCAATCAA	TGGCTACATA
3151	ATGGATACAC	TACCTGGCTT	AGTAATGGCT	CAGGATCAAA	GGATTGATG
3201	GTATCTGCTC	AGCATGGGCA	GCAATGAAAA	CATCCATTCT	ATTCAATTCA
3251	GTGGACATGT	GTTCACTGTA	CGAAAAAAAG	AGGAGTATAA	AATGGCACTG
3301	TACAATCTCT	ATCCAGGTGT	TTTGAGACA	GTGGAAATGT	TACCATCCAA
3351	AGCTGGAATT	TGGGGGTGG	AATGCCTTAT	TGGCGAGCAT	CTACATGCTG
3401	GGATGAGCAC	ACTTTTCTG	GTGTACAGCA	ATAAGTGTCA	GACTCCCTG
3451	GGAATGGCTT	CTGGACACAT	TAGAGATTTT	CAGATTACAG	CTTCAGGACA
3501	ATATGGACAG	TGGGCCCAA	AGCTGGCCAG	ACTTCATTAT	TCCGGATCAA
3551	TCAATGCCCTG	GAGCACCAAG	GAGCCCTTT	CTTGGATCAA	GGTGGATCTG
3601	TTGGCACCAA	TGATTATTCA	CGGCATCAAG	ACCCAGGGTG	CCCGTCAGAA
3651	GTTCTCCAGC	CTCTACATCT	CTCAGTTTAT	CATCATGTAT	AGTCTTGTATG
3701	GGAAGAAGTG	GCAGACTTAT	CGAGGAAATT	CCACTGGAAC	CTTAATGGTC
3751	TTCTTGCGCA	ATGTGGATT	ATCTGGATA	AAACACAATA	TTTTAACCC
3801	TCCAATTATT	GCTCGATACA	TCCGTTGCA	CCCAACTCAT	TATAGCATT
3851	GCAGCACTCT	TCGCATGGAT	TTGATGGGCT	GTGATTAAA	TAGTTGCAGC
3901	ATGCCATTGG	GAATGGAGAG	TAAAGCAATA	TCAGATGCAC	AGATTACTGC
3951	TTCATCCTAC	TTTACCAATA	TGTTGCCAC	CTGGTCTCCT	TCAAAAGCTC
4001	GAATTCACCT	CCAAGGGAGG	AGTAATGCCT	GGAGACCTCA	GGTGAATAAT
4051	CCAAAAGAGT	GGCTGCAAGT	GGACTTCCAG	AAGACAATGA	AAGTCACAGG
4101	AGTAACTACT	CAGGGAGTAA	AATCTCTGCT	TACCAGCATG	TATGTGAAGG
4151	AGTTCCCTCAT	CTCCAGCAGT	CAAGATGGCC	ATCAGTGGAC	TCTCTTTTT
4201	CAGAATGGCA	AAGTAAAGGT	TTTCAGGGAA	AATCAAGACT	CCTTCACACCC
4251	TGTGGTGAAC	TCTCTAGACC	CACCGTTACT	GACTCGCTAC	CTTCGAATT
4301	ACCCCCAGAG	TTGGGTGCAC	CAGATTGCC	TGAGGATGGA	GGTTCTGGGC
4351	TGCGAGGCAC	AGGACCTCTA	C		

FIG. 11B

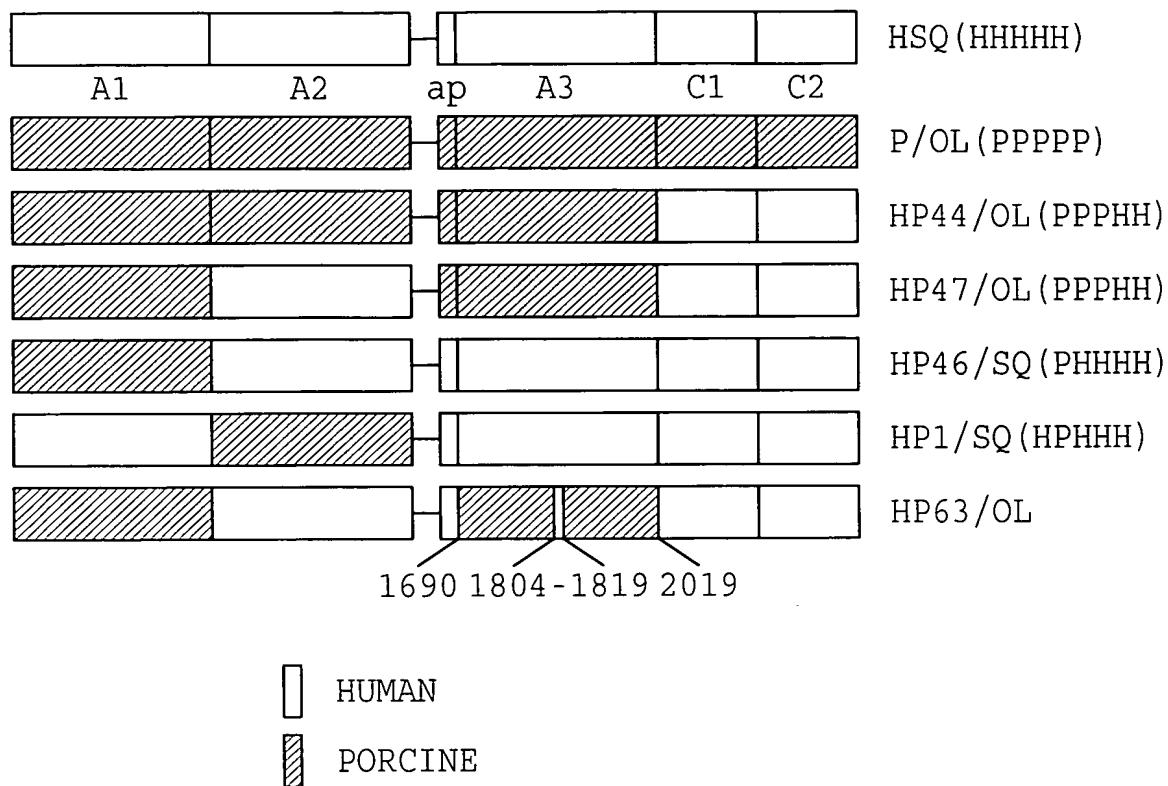


FIG. 12

AMINO ACID SEQUENCE OF HP630L

1	MQLELSTCVF	LCLLPLGFSA	IRRYYLGAVE	LSWDYRQSEL	LRELHVDTRF
51	PATAPGALPL	GPSVLYKKTV	FVEFTDQLFS	VARPRPPWMG	LLGPTIQAEV
101	YDTVVVTLKN	MASHPVSLHA	VGVSFWKSSE	GAEYEDHTSQ	REKEDDKVLP
151	GKSQTYWQV	LKENGPTASD	PPCLTYSYLS	HVDLVKDLNS	GLIGALLVCR
201	EGSLTRERTQ	NLHEFVLLFA	VFDEGKSWHS	ARNDSWTRAM	DPAPARAQPA
251	MHTVNGYVN	SLPGLIGCHK	KSVYWHVIGM	GTSPEVHSIF	LEGHTFLVRH
301	HRQASLEISP	LTFLTAQTFL	MDLGQFLLFC	HISSHGGM	EAHVRVESCA
351	EEPQLRRKAD	EEEDYDDNLY	DSDMDVVRLD	GDDVSPFIQI	RSVAKKHPKT
401	WVHYIAAEEE	DWDYAPLVLA	PDDRSYKSQY	LNNGPQRIGR	KYKKVRFMAY
451	TDETFKTREA	IQHESGILGP	LLYGEVGDTL	LIIFKNQASR	PYNIYPHGIT
501	DVRPLYSRRL	PKGVKHLKDF	PILPGEIFKY	KWTVTVEDGP	TKSDPRCLTR
551	YYSSFVNMER	DLASGLIGPL	LICYKESVDO	RGNQIMSDKR	NVILFSVFDE
601	NRSWYLTE	QRFLPNPAGV	QLEDPEFQAS	NIMHSINGYV	FDSLQLSVCL
651	HEVAYWYILS	IGAQTDLFSV	FFSGYTFKHK	MVYEDTLTLF	PFSGETVFMS
701	MENPGLWILG	CHNSDFRNRG	MTALLKVSSC	DKNTGDYYED	SYEDISAYLL
751	SKNNAIEPRS	FSQNSRHPST	RSQNPVVLKR	HQREITRTTL	QSDQEEIDYD
801	DTISVEMKKE	DFDIYDEDEN	QSPRSFQKRT	RHYFIAAVEQ	LWDYGMSESP
851	RALRNRAQNG	EVPRFKVVF	REFADGSFTQ	PSYRGELNKH	LGLLGPYIRA
901	EVEDNIMVT	KNQASRPYSF	YSSLISYPDD	QEQAEPKRN	FVKPNETKTY
951	FWKVQHHMAP	TEDEFDCKAW	AYFSDVDLEK	DVHSGLIGPL	LICRANTLNA
1001	AHGRQVTVQE	FALFFTIFDE	TKSWYFTENV	ERNCRAPCHL	QMEDPTLK
1051	YRFHÄINGYV	MDTLPGLVMA	QNQRIRWYLL	SMGSNENIHS	IHFSGHVFSV
1101	RKKEEYKMAV	YNLYPGVFET	VEMLPSKVGI	WRNRCLIGEH	LOAGMSTTFL
1151	VYSKKCQTPL	GMASGHIRDF	QITASGQYQGQ	WAPKLARLHY	SGSINAWSTK
1201	EPFSWIKVDL	LAPMIIHGIK	TQGARQKFSS	LYISQFIIMY	SLDGKKWQTY
1251	RGNSTGTLMV	FFGNVDSSGI	KHNI FN PPII	ARYIRLHPTH	YSIRSTLRME
1301	LMGCDLNSCS	MPLGMESKAI	SDAQITASSY	FTNMFATWSP	SKARLHLOGR
1351	SNAWRPQVNN	PKEWLQVDFQ	KTMKVTGVTT	QGVKSLLTS	YVKEFLISSS
1401	QDGHQWTLFF	QNGKVKVFQG	NQDSFTPVVN	SLDPPLLTRY	LRIHPQSWVH
1451	QIALRMEVLG	CEAQDLY			

FIG. 13

NUCLEOTIDE SEQUENCE OF HP63/OL

1	ATGCAGCTAG	AGCTCTCCAC	CTGTGTCTTT	CTGTGTCTCT	TGCCACTCGG
51	CTTTAGTGCC	ATCAGGAGAT	ACTACCTGGG	CGCAGTGGAA	CTGTCCTGGG
101	ACTACCAGCA	AAGTGAACTC	CTCCGTGAGC	TGCACGTGGA	CACCAAGATTT
151	CCTGCTACAG	CGCCAGGAGC	TCTTCCGTTG	GGCCCCTGAG	TCCTGTACAA
201	AAAGACTGTG	TTCGTAGAGT	TCACGGATCA	ACTTTTCAGC	GTTGCCAGGC
251	CCAGGCCACC	ATGGATGGGT	CTGCTGGGT	CTACCACATCA	GGCTGAGGTT
301	TACGACACGG	TGGTCGTTAC	CCTGAAGAAC	ATGGGCTTCTC	ATCCCCTTAG
351	TCTTCACGCT	GTCGGCGTCT	CCTTCTGGAA	ATCTTCCGAA	GGCGCTGAAT
401	ATGAGGATCA	CACCAAGCCAA	AGGGAGAAGG	AAGACGATAA	AGTCCTTCCC
451	GGTAAAAGCC	AAACCTACGT	CTGGCAGGTC	CTGAAAGAAA	ATGGTCCAAC
501	AGCCTCTGAC	CCACCATGTC	TTACCTACTC	ATACCTGTCT	CACGTGGACC
551	TGGTGAAAGA	CCTGAATTG	GGCCTCATTG	GAGCCCTGCT	GGTTTGTAGA
601	GAAGGGAGTC	TGACCAAGAGA	AAGGACCCAG	AACCTGCACG	AATTGTACT
651	ACTTTTGCT	GTCTTGATG	AAGGGAAAAG	TTGGCACTCA	GCAAGAAATG
701	ACTCCTGGAC	ACGGGCCATG	GATCCCGCAC	CTGCCAGGGC	CCAGCCTGCA
751	ATGCACACAG	TCAATGGCTA	TGTCAACACAGG	TCTCTGCCAG	GTCTGATCGG
801	ATGTCATAAAG	AAATCAGTCT	ACTGGCACGT	GATTGGAATG	GGCACCAAGCC
851	CGGAAGTGCA	CTCCATTGTT	CTTGAAGGCC	ACACGTTCT	CGTGAGGCAC
901	CATCGCCAGG	CTTCCTTGGA	GATCTCGCCA	CTAACTTTCC	TCACTGCTCA
951	GACATTCCCTG	ATGGACCTTG	GCCAGTTCT	ACTGTTTGT	CATATCTCTT
1001	CCCACCAACCA	TGGTGGCATG	GAGGCTCACG	TCAGAGTAGA	AAGCTGCC
1051	GAGGAGCCCC	AGCTCGGGAG	GAAAGCTGAT	GAAGAGGAAG	ATTATGATGA
1101	CAATTGTAC	GACTCGGACA	TGGACGTGGT	CCGGCTCGAT	GGTGACGACG
1151	TGTCTCCCTT	TATCCAAATC	CGCTCAGTTG	CCAAGAAGCA	TCCTAAAAT
1201	TGGGTACATT	ACATTGCTGC	TGAAGAGGAG	GACTGGGACT	ATGCTCCCTT
1251	AGTCCTCGCC	CCCGATGACA	GAAGTTATAA	AAGTCAATAT	TTGAACAATG
1301	GCCCTCAGCG	GATTGGTAGG	AAGTACAAAAA	AAGTCCGATT	TATGGCATAAC
1351	ACAGATGAAA	CCTTAAGAC	TCGTGAAGCT	ATTCAAGCATG	AATCAGGAAT
1401	CTTGGGACCT	TTACTTTATG	GGGAAGTTGG	AGACACACTG	TTGATTATAT
1451	TTAAGAATCA	AGCAAGCAGA	CCATATAACA	TCTACCCCTCA	CGGAATCACT
1501	GATGCCCGTC	CTTTGTATTG	AAGGAGATTA	CCAAAAGGTG	AAAACATTT
1551	GAAGGATTTT	CCAATTCTGC	CAGGAGAAAT	ATTCAAAATAT	AAATGGACAG
1601	TGACTGTAGA	AGATGGGCCA	ACTAAATCAG	ATCCTCGGTG	CCTGACCCGC
1651	TATTACTCTA	GTTCGTTAA	TATGGAGAGA	GATCTAGCTT	CAGGACTCAT
1701	TGGCCCTCTC	CTCATCTGCT	ACAAAGAAC	TGTAGATCAA	AGAGGAAACC
1751	AGATAATGTC	AGACAAGAGG	AATGTCATCC	TGTTTCTGT	ATTGATGAG
1801	AACCGAAGCT	GGTACCTCAC	AGAGAAATATA	CAACGCTTTC	TCCCCAATCC
1851	AGCTGGAGTG	CAGCTTGAGG	ATCCAGAGTT	CCAAGCCTCC	AACATCATGC
1901	ACAGCATCAA	TGGCTATGTT	TTTGATAGTT	TGCAGTTGTC	AGTTGTTTG
1951	CATGAGGTGG	CATACTGGTA	CATTCTAAC	ATTGGAGCAC	AGACTGACTT
2001	CCTTTCTGTC	TTCTCTCTG	GATATACCTT	CAAACACAAA	ATGGTCTATG
2051	AAGACACACT	CACCCATTTC	CCATTCTCAG	GAGAAACTGT	CTTCATGTCG
2101	ATGGAAAACC	CAGGTCTATG	GATTCTGGGG	TGCCACAACT	CAGACTTCG
2151	GAACAGAGGC	ATGACCGCCT	TACTGAAGGT	TTCTAGTTGT	GACAAGAAC
2201	CTGGTGATTA	TTACGAGGAC	AGTTATGAAG	ATATTCAGC	ATACTTGCTG
2251	AGTAAAACA	ATGCCATTGA	ACCTAGGAGC	TTCTCCAGA	ATTCAAGACA
2301	CCCTAGCACT	AGGTCTCAAA	ACCCACCACT	CTTGAAACGC	CATCAACGGG
2351	AAATAACTCG	TACTACTCTT	CAGTCAGATC	AAGAGGAAAT	TGACTATGAT

FIG. 14A

2401	GATACCATA	CAGTTGAAAT	GAAGAAGGAA	GATTTGACA	TTTATGATGA
2451	GGATGAAAAT	CAGAGCCCCC	GCAGCTTC	AAAGAGAAC	CGACACTATT
2501	TCATTGCTGC	GGTGGAGCAG	CTCTGGGATT	ACGGGATGAG	CGAATCCCC
2551	CGGGCGCTAA	GAAACAGGGC	TCAGAACGGA	GAGGTGCCTC	GGTTCAAGAA
2601	GGTGGTCTTC	CGGGAATTG	CTGACGGCTC	CTTCACGCAG	CCGTCGTACC
2651	GCGGGAACT	CAACAAACAC	TTGGGGCTCT	TGGGACCCCTA	CATCAGAGCG
2701	GAAGTTGAAG	ACAACATCAT	GGTAACCTTC	AAAAACCAAG	CGTCTCGTCC
2751	CTATTCCCTTC	TACTCGAGCC	TTATTCTTA	TCCGGATGAT	CAGGAGCAAG
2801	GGGCAGAAC	TCGAAAAAAAC	TTTGTCAAGC	CTAATGAAAC	CAAAACTTAC
2851	TTTTGGAAAC	TGCAGCATCA	CATGGCACCC	ACAGAAGACG	AGTTTGACTG
2901	CAAAGCCTGG	GCCTACTTTT	CTGATGTTGA	CCTGGAAAAA	GATGTGCACT
2951	CAGGCTTGAT	CGGCCCCCTT	CTGATCTGCC	GCGCCAACAC	CCTGAACGCT
3001	GCTCACGGTA	GACAAGTGAC	CGTGCAAGAA	TTTGCTCTGT	TTTCACTAT
3051	TTTTGATGAG	ACAAAGAGCT	GGTACTTCAC	TGAAAATGTG	GAAAGGAAC
3101	GCCGGGCC	CTGCCATCTG	CAGATGGAGG	ACCCCACACT	GAAAGAAAAC
3151	TATCGCTTCC	ATGCAATCAA	TGGCTATGTG	ATGGATACAC	TCCCTGGCTT
3201	AGTAATGGCT	CAGAATCAA	GGATCCGATG	GTATCTGCTC	AGCATGGGCA
3251	GCAATGAAA	TATCCATTG	ATTCAATTAA	GCGGACACGT	GTTCACTGTA
3301	CGGAAAAGG	AGGAGTATAA	AATGGCCGTG	TACAATCTCT	ATCCGGGTGT
3351	CTTGAGACA	GTGAAATGC	TACCGTCAA	AGTTGGAATT	TGGCGGAATA
3401	GATGCCGTAT	TGGCGAGCAC	CTGCAAGCTG	GGATGAGCAC	GACTTCCCTG
3451	GTGTACAGCA	AGAAGTGTCA	GACTCCCCTG	GGAATGGCTT	CTGGACACAT
3501	TAGAGATTTT	CAGATTACAG	CTTCAGGACA	ATATGGACAG	TGGGCCCAA
3551	AGCTGGCCAG	ACTTCATTAT	TCCGGATCAA	TCAATGCCTG	GAGCACCAAG
3601	GAGCCCTTTT	CTTGGATCAA	GGTGGATCTG	TTGGCACCAA	TGATTATTCA
3651	CGGCATCAAG	ACCCAGGGTG	CCCGTCAGAA	GTTCTCCAGC	CTCTACATCT
3701	CTCAGTTTAT	CATCATGTAT	AGTCTTGATG	GGAAGAAGTG	GCAGACTTAT
3751	CGAGGAAATT	CCACTGGAAC	CTTAATGGTC	TTCTTGGCA	ATGTGGATTC
3801	ATCTGGGATA	AAACACAATA	TTTTAACCC	TCCAATTATT	GCTCGATACA
3851	TCCGTTTGCA	CCCAACTCAT	TATAGCATTG	GCAGCACTCT	TCGCATGGAG
3901	TTGATGGGCT	GTGATTAAA	TAGTTGCAGC	ATGCCATTGG	GAATGGAGAG
3951	TAAAGCAATA	TCAGATGCAC	AGATTACTGC	TTCATCCTAC	TTTACCAATA
4001	TGTTGCCAC	CTGGTCTCCT	TCAAAAGCTC	GACTTCACCT	CCAAGGGAGG
4051	AGTAATGCCT	GGAGACCTCA	GGTGAATAAT	CCAAAAGAGT	GGCTGCAAGT
4101	GGACTTCCAG	AAGACAATGA	AAGTCACAGG	AGTAACTACT	CAGGGAGTAA
4151	AATCTCTGCT	TACCAGCATG	TATGTGAAGG	AGTTCCCTCAT	CTCCAGCAGT
4201	CAAGATGGCC	ATCAGTGGAC	TCTCTTTTT	CAGAATGGCA	AAGTAAAGGT
4251	TTTCAGGGA	AATCAAGACT	CCTTCACACC	TGTGGTGAAC	TCTCTAGACC
4301	CACCGTTACT	GACTCGCTAC	CTTCGAATTG	ACCCCCAGAG	TTGGGTGCAC
4351	CAGATTGCC	TGAGGATGGA	GGTTCTGGC	TGCGAGGCAC	AGGACCTCTA
4401	C				

FIG. 14B